ESA/P/WP.63

ENGLISH ONLY

January 1979

# MODEL LIFE TABLES FOR DEVELOPING COUNTRIES:

# AN INTERIM REPORT

prepared by

The Population Division, Department of International Economic and Social Affairs of the United Nations Secretariat.

79-42310

# I. INTRODUCTION

This paper is an interim report from a project designed to develop a new set of model life-tables. The distinguishing feature of the project is the attempt to make more extensive use of data from developing countries. Existing model life-table systems are almost entirely based on data from developed countries. There is mounting evidence that the age-sex pattern of mortality experienced in developing countries differs systematically, though not universally, from that in developed countries at equivalent mortality levels. In particular there is evidence of differing patterns in early childhood, with mortality between 6 months and two years or in age-group 1-4 often being relatively higher than in infancy (see e.g. Gordon, Wyon and Ascoli (1967); Dyson (1977); Simmons et al. (1978)). This phenomenon of relatively high mortality in early childhood has often been linked to weaning: the cessation of breast-feeding leading to exposure to micro-organisms and thus to diarrhoea and gastroenteritis (see e.g. Cantrelle and Leridon (1971)). In addition to patterns within childhood, the high incidence of diarrhoea and gastro-enteritis in tropical countries tends to bring about fairly extreme relationships between childhood mortality and adult mortality (see e.g. Sullivan (1973) and Cantrelle (1976)). These patterns do seem more extreme than are typically observed in developed countries, but they are more closely in line with the evidence of relatively high mortality under age 5 in Southern and Eastern Europe (see Coale and Demeny (1966) and Sullivan (1973)).

Although there is fairly considerable reason to believe that these pattern differences do exist, detailed information on mortality under age 5 and even within the first year of life is not commonly available for developing countries. Still less frequently available is information on weaning patterns. Even where it is available such information is rarely part of an overall study of mortality at all ages. Thus, any practical attempt to use information on mortality from developing countries at this time cannot fully take such factors into account, although their effects will clearly influence mortality estimates.

The availability of mortality data and estimates has certainly improved in the last ten or twenty years for developing countries. However, many of the available estimates are heavily based on existing model life-table systems. The degree of such dependence is variable. It is clearly unacceptable to use

data based upon an existing model life-table system as input to a new system. Thus, many available estimates, despite often being the best available for a particular society, are unusable for present purposes. Many of the better available estimates of mortality are based on questions about the survival of various relatives, especially of children, but also in decreasing order of frequency of parents, spouses and siblings. The techniques for converting the reported proportions surviving (or dead) into life-table measures all rely on existing model life-tables, at least to some extent. Provided that such estimates are taken only as indicators of mortality level and that little or no attempt is made to make inferences about pattern, they can be used to check other sources and even to provide a means of correcting other sources. Even though the estimates obtained may not be too heavily dependent on existing model life-tables, caution in their use is necessary and it is never possible to derive a full life-table from relative survival questions alone.

To obtain full life-tables for developing countries necessitates the availability of some sort of registration and population data or of survey data. It was for this reason that the extensive data gathering operations were carried out by OECD in cooperation with the U.N., as described in Section II. The availability of registration (or quasi-registration) and population data is however only the first step. There is every reason to suppose that such data will be deficient for most developing societies and, consequently, they cannot be used without evaluation and, if necessary, adjustment. Section III addresses the problems of making such evaluations and adjustments and gives a detailed illustrative example of the application of the procedures adopted to El Salvador. In view of the large number of data sets which required analysis within the evaluation and adjustment stage, these procedures had to be routinised and applied through a computer program. The volume of data precluded the use of one or two of the more ad hoc techniques which have been occasionally employed, although the results of such work are being utilised for those countries where they are usefully available, such as Colombia, Costa Rica and Korea. The adopted procedures are in process of being applied to all the data sources listed at Table II.2, and adjusted life-tables will be produced for those countries where the data prove useable.

After the derivation of adjusted life-tables for the time-points and countries available, it is necessary to combine the tables into a new set of model life-tables. Section IV considers this stage of the analysis, reporting on methodological

work to examine the reliability and efficiency of the various ways of generating model life-tables from observed (adjusted) data. This work was carried out by M.H. Le Bras of INED in consultation with the Population Division.

## II. SOURCES OF DATA

In this section, the main data sources which are being utilized in this project are described. In particular, it outlines the results of a project carried out by OECD in conjunction with the U.N. to gather together data on population and deaths by age and sex from developing countries and the subsequent adaptation of the data for further analysis.

The detailed sources for the data collected together by OECD are being published in three volumes entitled Mortality Project: Annotated Bibliography on the Sources of Demographic Data, OECD (1979). These volumes will be an outstanding bibliographic source of materials for the study of mortality data, giving information on the levels of detail available within each of the sources. The sources listed are not just those which provided the data base being utilised in this model life-table project. However, detailed information on these sources is not provided in the present document in view of the fairly complete documentation which is in course of publication by OECD.

For the model life-table project, the main value of the OECD work is the actual data brought together and partially processed. For as many countries as possible, both census and vital registration data were collected. In addition, with variable success by continent, an attempt was made to include surveys which included population and death data (by age and sex). Where possible, attempts were made to include data for urban and rural areas and native and foreign-born as well as for whole countries. Little or no attempt was made to collect sub-national data in most instances, but again this varied with continent.

The preparation of these data in machine-readable form and the preliminary analyses were carried out at the University of Louvain, Belgium with cooperation from OECD. For each country, many of which only had population data available, a variety of analyses were performed where applicable:

a) For each census: (i) a presentation of the numbers for the most detailed age-groups available

(ii) Age-sex pyramids and proportional distributions for 1 year, 5 year and 10-year age groups (where possible)

(iii) Calculation of several indicators of data reliability: sex ratios by age, Whipple's, Myer's and U.N. age-sex indices (see Section III.1), where possible.

b) For each year's registered deaths: sex ratios

c) Where both death registration and census data are available for some point in time:

(i) Calculation of death rates by age and sex ratios therefrom.

(ii) Application of Brass growth-balance method (see Section III.3) to age-range 10 to 64 using an ordinary least squares fit plus a graphical presentation.

(iii) Calculation of an unadjusted abridged life table giving ${}_nq_x$, $l_x$ and ${}_nd_x$.

(iv) Production of summary tables showing an index of the number of death at each age (and above each age) in terms of a base year.

These calculations and presentations were also performed (as far as possible) on data for subsets of the populations or on survey data. Although (a) and (b) are useful for assessing data quality in a broad sense, the most important input to the model life-table project comes from (c). Death rates were calculated for three-or-four-year periods centered on a census. For a census in January to March in year x, the deaths for the period x-2 to x+1 were averaged; for a census in April to September, deaths for the period x-1 to x+1 were averaged; and for a census in October to December, deaths for the period x-1 to x+2 were averaged.

In the OECD source, calculations of rates and life-tables were terminated with 75 and over as the open-ended group, even where data were available for higher ages. To supplement this source, detailed information was extracted from the raw data where possible to raise the terminal age-group to 85 and over. Furthermore, where possible we have supplemented the data source by adding deaths by single year of age under 5 from W.H.O. files. These files are another useful source of computer-based mortality information, being stored in more accessible form than the OECD files and thus allowing direct processing by the user. A problem with the W.H.O. files

is the lack of documentation and the lack of detail in the population estimates. Care was taken when mixing information from the OECD and WHO files to check that the two data sources were consistent. In a few instances, data were taken from Preston et. al. (1972) on deaths and population.

It was essential that most of the analyses (see Section III) be carried out on a computer, which necessitated routinely applicable techniques, and it was desired to handle as much of the available data as possible at the preliminary analysis and evaluation stage. As a result, the range of formats was restricted (in terms of age-groups) to a limited number, which are shown at Table II.1. These input formats are slightly adapted from those of the WHO data files, on which the program was initially tested. The data sets differ in the level of detail available for early childhood and at the older ages. The program is designed to handle any of these formats and automatically to align population and deaths with differing formats or two successive sets of data for a single country at differing time periods. As a result of the need to extend the level of age detail in several cases and the lack of an easily accessible, well-ordered file which was machine-readable, it was necessary to extract all the population and death information to be used as inputs to the evaluative stage and put them in the standardized formats for subsequent computer analysis.

Table II.2 shows in summary form the data available on population and deaths at this preliminary stage. Unless otherwise indicated, these data are based on census and registration data. It also shows the level of detail available (the format number), the source and the WHO country code (although in some instances these have been modified to deal with multiple surveys). At this preliminary stage there are available mortality and population data from 57 developing countries, or about 129 observed life tables (avoiding double counting for Afghanistan). As the project proceeds further life-tables may be added. Many of the life-tables included in Table II.2 will prove to be unsuitable for further study after the preliminary evaluation stage. For some of the surveys, the number of cases is too small for rates to be reliable. For some of the countries age-mistatement problems and under-registration are so severe as to preclude adequate adjustment. Furthermore, it is strongly preferable to use tables mainly (or even exclusively) from countries which have independent information from other sources against which to check levels of mortality. In particular, use of observed levels of childhood mortality without cross-checks against other sources causes concern (see Section III.5). At this stage in the project, such external information has not been systematically collated, but this is being done on a country by country basis as the evaluation proceeds. For each evaluation and adjustment of a particular country sources will be cited.

Table II.1 Formats used for input data to evaluation and analysis program

| Age group | Format number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| All ages | x | x | x | x | x | x | x | x |
| 0 | x | x | x | x | x | x | | x |
| 1 | x | | x | | | | | x |
| 2 | x | 1-4 | x | 1-4 | 1-4 | 1-4 | 0-4 | x |
| 3 | x | | x | | | | | x |
| 4 | x | | x | | | | | x |
| 5-9 | x | x | x | x | x | x | x | x |
| 10-14 | x | x | x | x | x | x | x | x |
| 15-19 | x | x | x | x | x | x | x | x |
| 20-24 | x | x | x | x | x | x | x | x |
| 25-29 | x | x | x | x | x | x | x | x |
| 30-34 | x | x | x | x | x | x | x | x |
| 35-39 | x | x | x | x | x | x | x | x |
| 40-44 | x | x | x | x | x | x | x | x |
| 45-49 | x | x | x | x | x | x | x | x |
| 50-54 | x | x | x | x | x | x | x | x |
| 55-59 | x | x | x | x | x | x | x | x |
| 60-64 | x | x | x | x | x | x | x | x |
| 65-69 | x | x | x | x | x | | x | |
| 70-74 | x | x | x | x | | | | |
| 75-79 | x | x | | | | 65+ | | 65+ |
| 80-84 | x | x | 75+ | 75+ | 70+ | | 70+ | |
| 85+ | x | x | | | | | | |
| Unknown | x | x | x | x | x | x | x | x |

x = data available for age group specified.

Table II.2 Sources of Life Tables

| Country | Code | Year | Format | Source | Notes |
|---|---|---|---|---|---|
| AFRICA | | | | | |
| Burundi | 1040 | 65 | 5 | OECD | Survey |
| | | 70 | 5 | " | Survey |
| West Cameroun | 1050 | 64 | 4 | OECD | Survey |
| | 1051 | 65 | 4 | " | Survey (Bamileke) |
| | 1052 | 65 | 4 | " | Survey (Environs of Bamileke) |
| Central African Empire | 1070 | 60 | 5 | OECD | Survey |
| Egypt | 1125 | 47 | 4 | OECD | |
| Guinea | 1130 | 54 | 4 | OECD | Survey |
| | 1131 | 57 | 4 | OECD | Survey (Konkoure) |
| Ethiopia | 1140 | 70 | 4 | OECD | Rural Survey |
| Kenya | 1220 | 68/70 | 2 | OECD | |
| Liberia (1240) | 1241 | 71 | 6 | OECD | Rural Survey |
| | 1242 | 71 | 6 | " | Urban Survey |
| Madagascar | 1260 | 66 | 5 | OECD | Survey |
| Mauritius | 1300 | 51/3 | 1 | OECD | |
| | | 61/3 | 1 | " | |
| | | 71/3 | 1 | " | |
| Morocco (1310) | 1311 | 72 | 4 | OECD | Survey (CERED) |
| Reunion | 1360 | 60/63 | 2 | OECD | |
| | | 66/69 | 4 | " | |
| Senegal (1390) | 1391 | 6/64 | 5 | OECD | Survey (Paoskoto) |
| | 1392 | 6/64 | 5 | " | Survey (Niakhar) |
| Seychelles | 1400 | 60 | 4 | OECD | Survey |
| Tanzania (1490) | 1491 | 67 | 4 | OECD | Rural Survey |
| | | 67 | 4 | " | Urban Survey |

Table II.2 Sources of Life Tables

| Country | Code | Year | Format | Source | Notes |
|---|---|---|---|---|---|
| AFRICA (continued) | | | | | |
| Tunisia | 1520 | 66 | 4 | OECD | Survey |
| | 1521 | 66 | 4 | " | (Rural) |
| | 1522 | 66 | 4 | " | (Urban) |
| | 1523 | 69 | 5 | " | Survey (Rural) |
| | 1524 | 69 | 5 | " | Survey (Rural) |
| Upper Volta | 1550 | 60 | 5 | OECD | Survey |
| LATIN AMERICA AND CARIBBEAN | | | | | |
| Argentina | 2020 | 59/61 | 5 | OECD | |
| | | 70 | 1 | WHO | |
| Barbados | 2040 | 59/61 | 2 | OECD | |
| | | 69/71 | 1 | " | |
| Belize | 2045 | 59/61 | 4 | OECD | |
| | | 69/71 | 4 | " | |
| Chile | 2120 | 09 | 2 | PKS | |
| | | 20 | 2 | " | |
| | | 29/32 | 2 | OECD | |
| | | 39/42 | 2 | " | |
| | | 51/53 | 2 | " | |
| | | 59/63 | 1 | OE+WH | |
| | | 69/71 | 2 | OECD | |
| Colombia | 2130 | 50/2 | 2 | OECD | |
| | | 63/5 | 2 | " | |
| Costa Rica | 2140 | 50 | 2 | OECD | |
| | | 62/4 | 2 | " | |
| | | 72/4 | 2 | " | |
| Dominican Republic | 2170 | 49/51 | 4 | OECD | |
| | | 59/61 | 4 | " | |
| | | 68/71 | 4 | " | |
| Ecuador | 2180 | 74 | 4 | OECD | |
| El Salvador | 2190 | 49/51 | 2 | OECD | |
| | | 60/2 | 1 | OE+WH | |
| | | 70/2 | 1 | " | |

Table II.2 Sources of Life Tables

| Country | Code | Year | Format | Source | Notes |
|---|---|---|---|---|---|
| LATIN AMERICA AND CARIBBEAN (continued) | | | | | |
| French Guinea | 2210 | 60/63 | 2 | OECD | |
| Guatemala | 2250 | 63/5 | 1 | OE+WH | |
| Guyana | 2260 | 59/61 | 2 | OECD | |
| Honduras | 2280 | 60/2 | 2 | OECD | |
| | | 71 | 1 | EDENH-Survey | |
| Mexico | 2310 | 38/41 | 2 | OECD | |
| | | 59/61 | 1 | OE+WH | |
| | | 68/71 | 1 | OE+WH | |
| Nicaragua | 2340 | 39/41 | 2 | OECD | |
| | | 50 | 2 | " | |
| | | 62/4 | 2 | " | |
| Panama | 2350 | 59/62 | 1 | OE+WH | |
| | | 69/71 | 1 | " | |
| Paraguay | 2360 | 71/3 | 4 | OECD | |
| Peru | 2370 | 72 | 2 | WHO | |
| Puerto Rico | 2380 | 49/51 | 2 | OECD | |
| | | 59/61 | 2 | OECD | |
| | | 69/71 | 2 | OE+WH | |
| Surinam | 2430 | 63/64 | 2 | OECD | |
| Trinidad and Tobago | 2440 | 45/7 | 2 | OECD | |
| | | 59/61 | 2 | " | |
| Uruguay | 2460 | 62/5 | 1 | OE+WH | |
| Venezuela | 2470 | 59/62 | 1 | OE+WH | |
| | | 70/3 | 2 | " | |
| ASIA | | | | | |
| Afghanistan | 3010 | 73 | 4 | OECD | Total Survey |
| | 3011 | 73 | 4 | " | Rural Survey |
| | 3012 | 73 | 4 | " | Urban Survey |
| | 3013 | 73 | 4 | " | Total Survey Adjusted |
| | 3014 | 73 | 4 | " | Rural Survey Adjusted |
| | 3015 | 73 | 4 | " | Urban Survey Adjusted |

Table II.2 Sources of Life Tables

| Country | Code | Year | Format | Source | Notes |
|---|---|---|---|---|---|
| ASIA (continued) | | | | | |
| Bangladesh (3025) | 3026 | 62 | 4 | OECD | Survey |
| | 3027 | 63 | 2 | " | Survey |
| | 3028 | 65 | 2 | " | Survey |
| | 3029 | 67 | 4 | " | Survey |
| | 3025 | 74 | 8 | MTLB | Survey-MatLab Thana |
| | | 75 | 8 | " | " |
| | | 76 | 8 | " | " |
| China (Taiwan) | 3070 | 20 | 2 | PKS | |
| | | 30 | 2 | " | |
| | | 39/42 | 2 | OECD | |
| | | 55/57 | 4 | " | |
| | | 70/71 | 4 | " | |
| Hong Kong | 3090 | 59/62 | 4 | OECD | |
| | | 69/72 | 1 | CE+WH | |
| Iran | 3130 | 75 | 2 | OECD | Survey |
| Israel | 3150 | 48 | 2 | OECD | Jewish population |
| | | 64 | 2 | PKS | " |
| | | 72 | 4 | OECD | " |
| Jordan | 3170 | 61/62 | 4 | OECD | Survey |
| | 3171 | 61 | 4 | " | " |
| | 3172 | 72 | 5 | " | " |
| | 3173 | 75 | 6 | " | " |
| Korea | 3180 | 66/67 | 2 | OECD | |
| Kuwait | 3190 | 70 | 6 | OECD | |
| | | 75 | 4 | " | |
| Sabah | 3240 | 59/61 | 4 | OECD | |
| | | 69/71 | 6 | " | |
| Malaysia | 3252 | 56/58 | 5 | OECD | |
| | | 69/71 | 6 | " | |
| | 3253 | 69/71 | 6 | " | (Rural) |
| | 3254 | 69/71 | 6 | " | (Urban) |

Table II.2 Sources of Life Tables

| Country | Code | Year | Format | Source | Notes |
|---|---|---|---|---|---|
| ASIA (continued) | | | | | |
| Pakistan | 3290 | 63 | 2 | OECD | Survey |
| | 3291 | 64 | 2 | " | " |
| | 3292 | 65 | 2 | " | " |
| | 3293 | 68 | 6 | " | " |
| | 3294 | 68 | 6 | " | Survey (Rural) |
| | 3295 | 69 | 6 | " | Survey (Urban) |
| Philippines | 3300 | 58/61 | 4 | OECD | |
| | | 69/71 | 4 | " | |
| Singapore | 3350 | 56/58 | 2 | OECD | |
| | | 69/71 | 4 | " | |
| Sri-Lanka | 3365 | 51/54 | 2 | OECD | |
| | | 60 | 2 | PKS | |
| Thailand | 3380 | 59/61 | 2 | OECD | |
| | | 69/71 | 2 | " | |
| | 3381 | 65 | 6 | " | (survey) |

SOURCES: OECD - Organisation for Economic Cooperation and Development - Mortality Project.
WHO - World Health Organisation standard data tapes.
PKS - Preston, Keyfitz and Schoen.
MTLB - MatLab project, Cholera Research Laboratory, Dacca, Bangladesh.
EDENH - Encuesta Demografía Nacional de Honduras.
OEWH - combined from OECD and WHO.

## III. EVALUATION OF AVAILABLE DATA

In the previous section the major data sources used in this project were described, in particular the cooperative project with OECD which gathered together registration and quasi-registration data from the third world. The major problem with these data is one of quality: age-mistatement is a well recognised problem in developing countries; it is important to try to assess whether all deaths are registered and whether omissions are selective by age or sex; censuses also present problems of underenumeration, which may be selective. These areas are now considered.

There are two main approaches to the evaluation of demographic data, namely checking internal consistency and checking against external data sources. Both are useful and ideally prove mutually reinforcing. Many such approaches throw light on inadequacies of the data, but it is much harder to produce techniques which provide a means of adjusting for the deficiencies so detected. Most such corrective techniques rely on the use of model life-tables, which is clearly not acceptable to provide input to a project to produce further model life-tables. Nevertheless it is clear that corrective techniques are vital in this work if maximal use is to be made of available data sources.

Each of the evaluation and correction techniques tried in the project will first be reviewed, with illustrative examples, and then their relative merits will be discussed. Not all the techniques have proved successful, but it is important to review this experience.

### 1. Indices of age-mistatement.

These indices are generally treated as being non-corrective and are based on internal consistency only. As part of the OECD input to the project, Myers', Whipple's and U.N. age-sex indices were calculated for the population data. As these indices are quite well-known, a complete description is not given here, but the interested reader is referred to U.N. Manual II (1956) or to chapter 8 of Shryock and Siegel (1971). At best these indices give a broad indication of data quality, but only for the population data, not for the deaths or the death rates. It is probable that poor data on ages for the population will be associated with similar poor quality for the deaths.

Table III.1. Indices of Age-Mis-statement

| Country | Census date | Both Sexes | | U.N. | | |
|---|---|---|---|---|---|---|
| | | Whipple | Myers | Age regularity | Sex regularity | Compound |
| Argentina | 9/70 | 103.89 | 2.20 | 2.87 | 2.20 | 12.41 |
| Argentina | 9/60 | 104.78 | 2.60 | 3.14 | 2.24 | 13.21 |
| Uruguay | 10/63 | 105.83 | 4.10 | 3.65 | 2.94 | 16.35 |
| Venezuela | 2/61 | 120.26 | 7.30 | 4.72 | 4.05 | 21.62 |
| Chile | 5/52 | 128.39 | 14.30 | 8.64 | 4.66 | 31.72 |
| Chile | 11/60 | 129.84 | 12.20 | 4.11 | 2.97 | 17.22 |
| Colombia | 7/64 | 143.49 | 16.00 | 9.36 | 4.26 | 31.81 |

Illustrations of these indices are presented at Table III.1. Whipple's index has a lower bound of approximately 100 and could reach 500 if everyone gave ages which were multiples of 5. It is an index of the relative preference for ages ending in zero and five, with a value around 100 representing no preference. Myers' index includes some allowance for preference for ages other than multiples of five and is based on estimates of the relative attraction of each terminal digit. Its value has a lower limit of zero and an upper limit of 90 if everyone preferred a single terminal digit. The U.N. age-and sex- regularity and compound indices are based solely on five-year age-groups and thus give an indication of major distortions in the age-sex distributions, although some deviations from a regular declining age-distribution are likely to be genuine reflections of past demographic events.

All these indices give at least an indication of data quality for the age-reckoning of the census (or other) sources of population data. They give no direct information on the age-mis-statement patterns for the deaths. Neither do such indices provide any information on underenumeration, whether age-selective or not. For what it is worth the illustrative examples suggest age data in Argentina to be quite good, in Colombia to be more suspect. In general these indices have not proved very useful for this project.

2. Sex ratios

A simple, but sometimes powerful approach to examining overall quality of deaths, population and death rates based on them is an examination of the sex-ratios of each. A major problem with this approach is that it requires the imposition of a normative pattern of sex-ratios. If these are derived from model life tables there would be inherent dangers in rejecting data as untrustworthy simply because of non-concurrence with existing model life-table patterns, expecially as the work linking the sexes in model life table systems has been perfunctory. An example of fairly well-established deviant patterns of sex ratios in mortality is given in a recent unpublished note by Coale and Goldman (1978) on patterns of mortality in the Far East, which shows extremely high sex ratios of mortality for the older ages to be a consistent feature of this geographical area. Thus, useful insights can be gained into the patterns of mortality and to a lesser

extent into the patterns of errors in the data from an examination of the sex ratios.

3. The Brass growth-balance method and related approaches

We turn now to a group of methods which are internal consistency checks but have an inbuilt corrective element (at least when applicable). These methods are recent, the first, informal, appearance having been in 1973. The best available accounts of the method are given in Brass (1975) or the forthcoming manual being produced by the Committee on Population and Demography of the American National Academy of Sciences. In essence this method makes use of the simple identity, $b(a+) = r(a+) + d(a+)$, which always holds for a closed population, with $r(a+)$ being the rate of increase of the population aged a and over, $d(a+)$ the death rate at ages a and over and $b(a+)$ the 'birth rate' at ages a and over, or the rate of entry into the population at age a. Brass' version of the technique made $r(a+)$ constant, which implies that the population can be taken to be stable. Such an assumption clearly limits the applicability of the technique in this form, as few countries can now be considered to have populations which are even approximately stable. Any application of the growth-balance method, or related approaches, has to make an additional simplifying assumption about the pattern of omission of deaths. This is almost always of the form that for a segment of the age-range (usually a and over) registration is deficient by a constant proportion at all ages within the segment. In addition, it is implicit that the estimates based on the age distribution (namely of $b(a+)$ and possibly $r(a+)$) are virtually correct, or at least do not suffer from systematic errors. Thus if the estimate of $d(a+)$, based on registered deaths, is denoted by $d^*(a+)$ and the correction factor required is denoted by f, $d(a+) = f\, d^*(a+)$ and $b(a+) = r(a+) + f\, d^*(a+)$ or for Brass' particularisation to a stable population $b(a+) = r + f\, d^*(a+)$. Provided that it is possible to successfully estimate $b(a+)$, $d(a+)$ (and $r(a+)$ when applicable) and to validly make the necessary assumptions of a constant proportion of deaths being unregistered throughout the age range a+ and of stability for Brass' approach, it is simple to use this technique. In practice there are several difficulties. There are almost always errors in the data which are not in conformity with the assumptions. In an attempt to overcome these problems Brass' original suggestion was to regress $d^*(a+)$ on $b(a+)$ for a range of values of a, thus 'averaging out'

the errors to some extent. This raises the problem of which range of values of a to use. It is usually assumed that registration in childhood may be differentially deficient, as is often the population enumerated in the youngest ages. It is also clear that overstatement of age becomes a major problem at the highest ages, not necessarily being of the same order for the population (many of whom report their own ages) and the deaths (none of whom report their own ages). Thus the age-range for fitting the regression line is usually restricted, often to something like ages 15 to 65. Some authors use the 'group-average' method (Bartlett (1949)), as it is reasonable to presume errors in both variables; this is sometimes used to fit 9 lines through all combinations of ages, from each of the starting ages 15, 20 and 25 to each of the finishing ages 55, 60 and 65. (See Potter and Ordóñez, (1976), Barclay et. al. (1976) and Martin (1977 and 1978)), taking the median value as the best estimate. We have preferred to use the biweight regression procedure due to Tukey (Mostellar and Tukey (1977) and McNeil (1977) from whence the computer program used was adapted), although conventional least-square estimation procedures have been tried as well. These regressions have been fitted over several different age-ranges.

One problem which often seems to occur with the Brass method is the regression line (however fitted) giving an apparently good fit and yet implying deaths to be substantially over-registered. This problem was first brought to attention by Potter and Ordóñez (1976) for the case of Colombia. Potter and Ordóñez adopted some quite clever adjustment procedures to try to overcome these problems. Their results on Colombia will be used as being the best available. However, it is not always possible to carry out adjustment procedures similar to those of Potter and Ordóñez, which require time-series data or fitting of stable populations. In some instances, then, it has been necessary to abandon or ignore the results of the Brass procedures for these reasons. Of course such findings also raise doubts about the results from the Brass procedure which do appear plausible. If, for example, death registration completeness of around 90% can give estimates

of 110% completeness, what does an estimated completeness of around 90% mean? If the same pattern of errors or biasses is present this could reflect only 70% completeness.

In addition Potter and Ordóñez reported some results on the impact of rapid mortality decline, as experienced in Colombia, on estimates from the Brass technique. They found that rapid mortality decline leads to underestimates of the completeness of registration.

In response to some of these problems which have arisen with the original Brass technique other approaches have been developed. The two utilized in this project are due to Martin (1977 and 1978) and Preston (1978). As these two approaches are very recent developments, the present project constitutes the first systematic use of these techniques, as well as the most extensive use of the original Brass method. Thus, results of these applications require careful evaluation; nevertheless, it was necessary to try all available techniques, especially those with a corrective element.

The first development from Brass' technique utilized here is the approach due to Martin (1977, 1978) which is adapted somewhat here. This is essentially a return to the more general relationship, $b(a+) = r(a+) = d(a+)$. The additional problem becomes the estimation of $r(a+)$. The approach used here involves taking intercensal growth rates (where possible) to obtain estimates of $r(a+)$, although Martin tries to avoid this in her work. Using intercensal growth rates obviously presents problems: the aim is to estimate completeness at the beginning or end of the period, yet use is made of an average growth rate for the period; and the estimates are vulnerable to differences in completeness of census enumeration. Where possible, both retrospective and prospective estimates of $r(a+)$ are used, but this requires three successive censuses. Average of the two sets of $r(a+)$ are not used, but rather two sets of estimates are obtained, since the last 20 or 30 years have seen such rapid demographic change in developing countries.

Once estimates of $r(a+)$ are available, the use of the procedure is a relatively simple matter. The growth balance equation is rearranged as $b(a+) - r(a+) = f \cdot d^*(a+)$ and, by regressing $d^*(a+)$ on $b(a+) - r(a+)$, an estimate of the completeness of death registration is obtained on the

usual assumption of constant proportions omitted by age, at least for ages a and above. These regression lines are fitted without forcing them through the origin, in the belief that the magnitude of the constant term (which should be zero) will help to provide evidence of failures of the technique or departures from the assumptions. Again a case could be made for the use of Bartlett's 'group-average' method for fitting the regressions as there are likely to be errors in the variables, although Durbin's (1954) method is generally held to be more efficient. Here use is again made of the biweight procedure due to Tukey which has the advantage over the other methods of letting the weights be determined by the data in a more flexible way. Little is known about the efficiency of the Tukey procedures.

In addition, the value of the constant term provides an estimate of differential enumeration between the two censuses under some simplifying assumptions. If the difference in completeness between the two censuses is constant with age above the lowest age used, then the value of the constant obtained leads directly to an estimate of this omission rate, being a measure of the proportional omission per year of the intercensal period. Inflation back to the relevant intercensal period gives an estimate of the relative underenumeration, which will refer to the earlier census if the constant is negative, for the later census if positive. Naturally this requires that the assumptions implicit in the technique (constant proportions of deaths missed with age as well as the constant proportional enumeration differences) with age be satisfied for the estimate to be meaningful. In view of some of the problems experienced with this technique, which will be described later, it is doubtful whether any such estimates of relative underenumeration should be taken too seriously.

Preston (1978) gives two further methods for estimating death registration completeness, the first being a substantial modification of Brass' growth balance technique and the second using intercensal cohort deaths and providing an estimate of relative incompleteness of the less complete of the two censuses as well as of death registration. Owing to pressures of time and the difficulties of obtaining complete series of intercensal

registered deaths for developing countries we have only used the first of these methods here. A systematic application of the second approach is a part of the work programme of the Population Division for 1979-81.

Preston's first approach involves the application of a formula rather than the fitting of a curve:

$$C = \frac{d^{*}_{x+}}{r} \left[ \int_{x}^{\infty} d^{1}_{x+}(a)\ e^{r(a-x)}\ da - 1 \right], \text{ where}$$

$C$ = proportionate completeness of death registration

$d^{1}_{x+}(a)$ = proportion of death above age x that occur at age a.

In practice an estimate of r is required as is an approximation to the integral by a discrete form

$$\sum_{x}^{w} d^{1}_{x+}\ (a \text{ to } a+n)\ e^{r(a+n/2 - x)}, \quad \text{taking the average}$$

age in each n-year group as an approximation to that implied by the mean value theorem. The estimation of r presents more of a problem, but can be taken from a survey or from intercensal rates of growth. When using intercensal rates of growth a case can be made for using the growth rate for ages x and over, especially in the somewhat unlikely situation of a fairly recent fertility decline being the sole reason for departures from stability. If the population were truly stable this would be unnecessary.

As the Brass approach could also be used to obtain a point estimate of completeness with a given value of r, using $k = \frac{b_{x+}}{d^{*}_{x+}} - r$, it is helpful

to contrast the two approaches to highlight the differences and the reasons for Preston's suggested alternative method. Both require estimates of $d^*_{x+}$ and the assumption of stability; both use an estimate of r. It is in the further information used that they differ. Brass' method requires an estimate of $b_{x+}$, which is the rate at which entrants pass age x and is usually estimated by $\frac{C(x+5) - C(x-5)}{10\ C(x)}$, where C(x) is the proportion of the population aged x and over. (This procedure is also used to estimate $b_{x+}$ for the Martin modification of Brass' technique). This estimation is generally held to be particularly vulnerable to age-mis-statement due to the use of a ten-year age group in the numerator, all the other measures being based on ages x and over and thus only sensitive to crossing of a single boundary at age x. In addition such simple averaging can be held to be suspect for point estimates. It was in an attempt to overcome this problem that Preston devised his alternative technique. Instead of estimating $b_{x+}$ it becomes necessary to estimate $\int_x^w d^1_{x+}(a)\ e^{r(a-x)}da$

or its discrete approximation $\sum_{a-x}^{w-5} \left[ d^1_{x+}(a \text{ to } a+5)\ e^{r(a-x+2.5)} \right]$.

As the weights (exponential terms) applied to the death distribution increase with increasing age, this technique is quite vulnerable to overstatement of age at death and thus, in turn, is suspect. Overstatement of age will lead to overestimation of the integral and thus overestimates completeness of registration C or underestimates the inflation factor k. On the other hand Brass' technique can be held to be insensitive to overstatement of age if the points used are only based on ages below those subject to overstatement.

## 4. Examples

While there is some advantage to providing carefully chosen illustrative examples from several countries it has been decided to present several examples from a single country to permit comparisons between the methods used. Somewhat, but not entirely, arbitrarily El Salvador was selected for these illustrations. Data are available at three points in time for

El Salvador, which allows comparison of techniques based on forward and backward intercensal growth rates for the central data point. Thus although data are available for the three periods, 1949/51, 60/62 and 70/72 attention is concentrated mainly on the analysis for the life table based on the deaths for 1960/62 and the census of May 1961. Illustrative calculations for females are shown first and then compared with the results for males and for the other periods. Such comparisons provide a powerful consistency check.

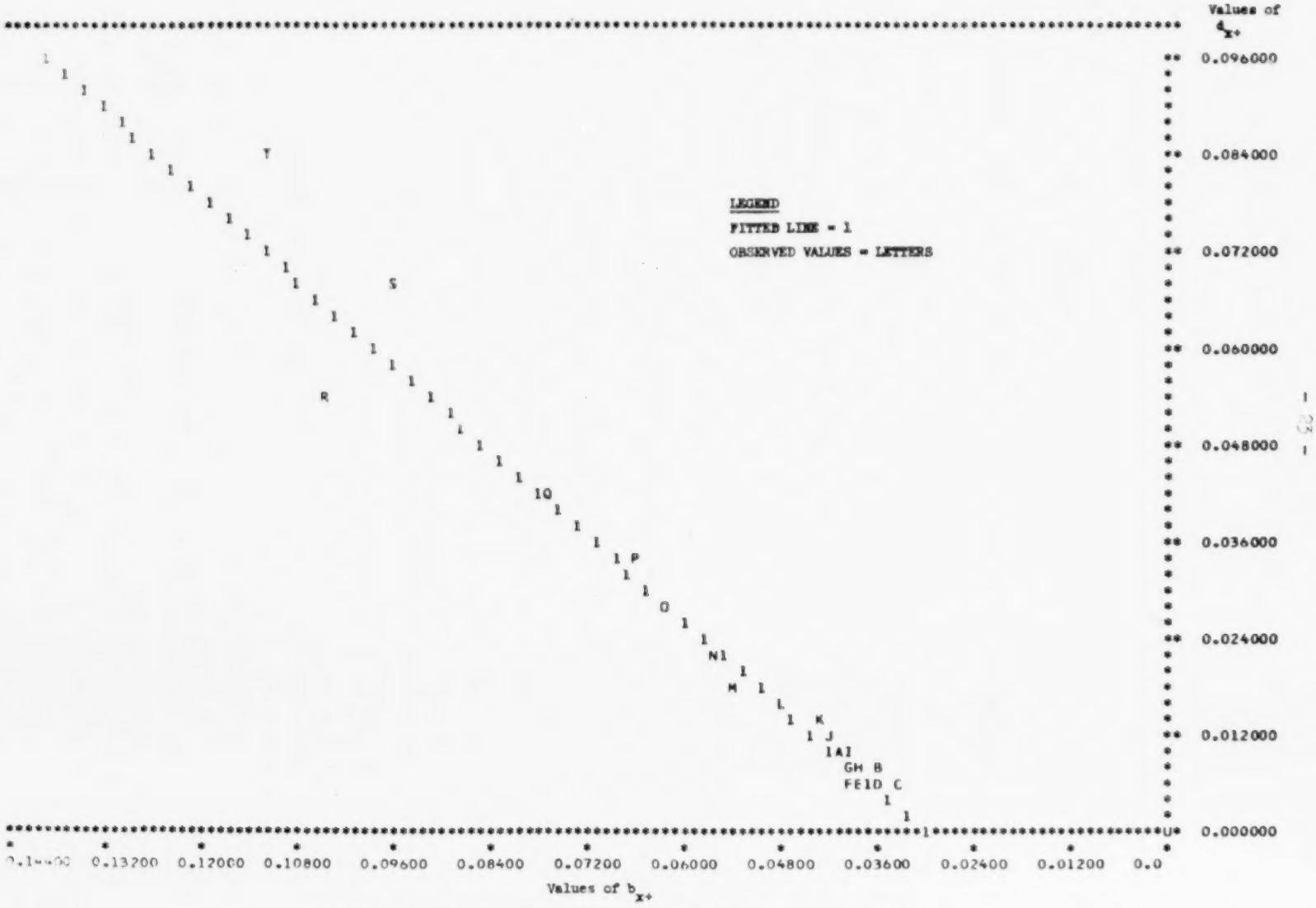
Tables III.2 to III.4 illustrate the application of Brass' original growth-balance technique, with D(Y)/P(Y) representing d(y+) and N(Y)/P(Y) representing b(y+). The fit shown in Table III.2 is that to ages 5+ to 60+ , using a biweight regression procedure. Table III.3 graphs the fit and the actual values to aid assessment of departures from linearity. The successive ages are denoted by A, B, C, ...., S, T. Table III.4 shows the results of several regression fits over different age ranges using both ordinary least squares and biweight regression procedures. From Table III.3 it is clear that there do exist distortions at the high ages, almost certainly due to overstatement of age, which in itself would lead to the observed pattern, but perhaps also due to even more overstatement of age for decedents. It is also clear from Table III.4 that the biweight procedure is not of itself able to eliminate the effects of these deviant points. What clearly happens is that any initial fit gets pulled close to ages 70 and 75 (points S and T), and subsequent interations give low weights to age 65 as being deviant. Thus, while the biweight procedures are relatively robust, they are not able to deal with a high proportion of inconsistent points. However it is also clear from Table III.4 that there is a high degree of consistency in the estimates of registration completeness based on the age points 1 through 60 or any of the subsets used, especially using the more robust biweight regression procedures. Unfortunately results from this technique are not always so consistent or plausible. As an example similar tables are presented for El Salvador females for the period centered on the 1971 census in Tables III.5 to III.7. Here the phenomenon to which Potter and Ordóñez first drew attention is apparent, there being apparent overregistration of deaths and the estimates being more variable. By 1971 the population of El Salvador was unlikely to have been even approximately

TABLE III.2. RESULTS OF BRASS GROWTH-BALANCE TECHNIQUE FOR EL SALVADOR, FEMALES, 1961.

| AGES | POPULATION | DEATHS | CUMULATED POPULATION | CUMULATED DEATHS | $d_{x+}$ | $b_{x+}$ | FIT * | RESIDUALS | UNADJUSTED RATES | ADJUSTED RATES | AGES |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 48114. | 4073. | 1274256. | 13654. | 0.0107153 | 0.0410122 | 0.0425334 | -0.0015213 | 0.0846530 | 0.0959295 | 0 |
| 1 | 39822. | 1349. | 1226142. | 9581. | 0.0078139 | 0.0358590 | 0.0392456 | -0.0033866 | 0.0338821 | 0.0383955 | 1 |
| 2 | 40183. | 719. | 1186320. | 8232. | 0.0069389 | 0.0337202 | 0.0382540 | -0.0045338 | 0.0178930 | 0.0202765 | 2 |
| 3 | 43330. | 482. | 1146136. | 7513. | 0.0065548 | 0.0365636 | 0.0378188 | -0.0012552 | 0.0110480 | 0.0125196 | 3 |
| 4 | 42340. | 286. | 1102506. | 7031. | 0.0063770 | 0.0389885 | 0.0376173 | 0.0013712 | 0.0067641 | 0.0076651 | 4 |
| 5 | 190234. | 610. | 1060166. | 6744. | 0.0063616 | 0.0392621 | 0.0379998 | 0.0016624 | 0.0032071 | 0.0036343 | 5 |
| 10 | 149538. | 214. | 869932. | 6134. | 0.0070514 | 0.0390574 | 0.0383815 | 0.0006759 | 0.0014296 | 0.0016201 | 10 |
| 15 | 125040. | 266. | 720394. | 5920. | 0.0082183 | 0.0381151 | 0.0397039 | -0.0015888 | 0.0021291 | 0.0024127 | 15 |
| 20 | 113490. | 291. | 595353. | 5654. | 0.0094972 | 0.0400653 | 0.0411531 | -0.0010878 | 0.0025679 | 0.0029100 | 20 |
| 25 | 91663. | 271. | 481863. | 5363. | 0.0111292 | 0.0425750 | 0.0430025 | -0.0004276 | 0.0029594 | 0.0033536 | 25 |
| 30 | 77711. | 315. | 390200. | 5092. | 0.0130485 | 0.0434071 | 0.0451774 | -0.0017703 | 0.0040487 | 0.0045880 | 30 |
| 35 | 72936. | 349. | 312489. | 4777. | 0.0152866 | 0.0482090 | 0.0477137 | 0.0004953 | 0.0047838 | 0.0054210 | 35 |
| 40 | 56942. | 338. | 239552. | 4428. | 0.0184843 | 0.0542171 | 0.0513374 | 0.0028797 | 0.0059327 | 0.0067230 | 40 |
| 45 | 46205. | 357. | 182610. | 4090. | 0.0223982 | 0.0564846 | 0.0557726 | 0.0007119 | 0.0077261 | 0.0087552 | 45 |
| 50 | 36616. | 385. | 136406. | 3733. | 0.0273681 | 0.0621828 | 0.0614046 | 0.0007782 | 0.0099755 | 0.0113044 | 50 |
| 55 | 26154. | 387. | 97790. | 3348. | 0.0342363 | 0.0662337 | 0.0691876 | -0.0029540 | 0.0148062 | 0.0167785 | 55 |
| 60 | 29273. | 647. | 71636. | 2961. | 0.0413300 | 0.0773726 | 0.0772263 | 0.0001463 | 0.0221160 | 0.0250620 | 60 |
| 65 | 14964. | 449. | 42363. | 2313. | 0.0546071 | 0.1044246 | 0.0922720 | 0.0121527 | 0.0299881 | 0.0339828 | 65 |
| 70 | 11705. | 504. | 27399. | 1865. | 0.0680530 | 0.0955135 | 0.1075090 | -0.0119955 | 0.0449972 | 0.0509912 | 70 |
| 75 | 6902. | 396. | 16193. | 1360. | 0.0840069 | 0.1118221 | 0.1255881 | -0.0137660 | 0.0574157 | 0.0650639 | 75 |
| UNKNOWN | ( 268.) | (114.) | | | | | | | | | |
| ALL | 1274256. | 13654. | | | | | | | 0.0107153 | 0.0121426 | ALL |

* Fit based on biweight regression through ages 5+ to 60+ giving r = 0.05039 and f = 1.13321.

TABLE III.3. GRAPH OF OBSERVED AND FITTED POINTS FOR BRASS GROWTH-BALANCE METHOD, EL SALVADOR, FEMALES, 1961.




BIWEIGHT REGRESSION ESTIMATES BASED ON AGES 5+ TO 60+

(R = 0.05039 FACTOR = 1.13321)

TABLE III.4. RESULTS OF BIWEIGHT AND LEAST SQUARES REGRESSIONS THROUGH VARIOUS AGE-GROUPS, BRASS GROWTH-BALANCE TECHNIQUE. EL SALVADOR, FEMALES, 1961 (SHOWING WEIGHTS USED IN REGRESSIONS AND RESIDUALS X 1000).

| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

BIWEIGHT REGRESSION ON AGE-GROUPS 0- TO 80
INTERCEPT(EST.OF R)= 0.031812 SLOPE (REGIST.FAC) = 0.977883 COMPLETENESS 1.023035
WEIGHT 0.990 0.926 0.865 0.984 0.995 0.991 0.999 0.982 0.994 1.000 0.992 0.988 0.893 0.955 0.925 0.995 0.850 0.0 0.954 0.974 0.0
RESIDS -1.27 -3.59 -4.87 -1.66 0.94 1.23 0.35 -1.73 -1.03 -0.12 -1.16 1.46 4.34 2.78 3.62 0.96 5.16 19.24 -2.82 -2.10 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 0- TO 80
INTERCEPT(EST.OF R)= 0.031304 SLOPE (REGIST.FAC) = 1.042425 COMPLETENESS 0.959302
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS -1.46 -3.59 -4.82 -1.58 1.04 1.32 0.40 -1.76 -1.14 -0.33 -1.50 0.97 3.64 1.83 2.35 -0.76 2.98 16.20 -6.73 -7.05 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 1- TO 80
INTERCEPT(EST.OF R)= 0.031927 SLOPE (REGIST.FAC) = 0.975893 COMPLETENESS 1.024703
WEIGHT 0.0 0.927 0.869 0.983 0.996 0.993 1.000 0.982 0.993 1.000 0.991 0.990 0.904 0.961 0.933 0.996 0.863 0.0 0.957 0.976 0.0
RESIDS****** -3.69 -4.98 -1.76 0.84 1.13 0.25 -1.83 -1.13 -0.21 -1.25 1.36 4.25 2.70 3.55 0.90 5.11 19.21 -2.83 -2.09 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 1- TO 80
INTERCEPT(EST.OF R)= 0.031430 SLOPE (REGIST.FAC) = 1.040440 COMPLETENESS 0.961132
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -3.70 -4.93 -1.69 0.92 1.21 0.29 -1.87 -1.25 -0.43 -1.60 0.87 3.56 1.75 2.28 -0.82 2.94 16.18 -6.72 -7.01 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 1- TO 60
INTERCEPT(EST.OF R)= 0.029832 SLOPE (REGIST.FAC) = 1.147805 COMPLETENESS 0.871228
WEIGHT 0.0 0.607 0.332 0.968 0.835 0.782 0.936 0.934 0.977 1.000 0.902 0.965 0.553 0.955 0.956 0.618 0.999 0.0 0.0 0.0 0.0
RESIDS****** -2.94 -4.08 -0.79 1.84 2.13 1.13 -1.15 -0.67 -0.03 -1.40 0.83 3.17 0.94 0.94 -2.89 0.10********** 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 1- TO 60
INTERCEPT(EST.OF R)= 0.029575 SLOPE (REGIST.FAC) = 1.152879 COMPLETENESS 0.867393
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -2.72 -3.85 -0.57 2.06 2.35 1.35 -0.94 -0.46 0.17 -1.21 1.01 3.33 1.09 1.06 -2.81 0.15********** 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 80
INTERCEPT(EST.OF R)= 0.033181 SLOPE (REGIST.FAC) = 0.955856 COMPLETENESS 1.046182
WEIGHT 0.0 0.0 0.0 0.0 0.0 1.000 0.997 0.962 0.979 0.993 0.978 0.999 0.950 0.984 0.964 1.000 0.905 0.038 0.969 0.988 0.0
RESIDS********** 0.00 -0.86 -2.92 -2.19 -1.24 -2.25 0.42 3.37 1.89 2.84 0.33 4.69 19.05 -2.72 -1.06 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 80
INTERCEPT(EST.OF R)= 0.032733 SLOPE (REGIST.FAC) = 1.016316 COMPLETENESS 0.983946
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********** 0.06 -0.84 -2.97 -2.32 -1.47 -2.59 -0.06 2.70 0.99 1.64 -1.29 2.64 16.19 -6.30 -6.29 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 60
INTERCEPT(EST.OF R)= 0.030391 SLOPE (REGIST.FAC) = 1.133208 COMPLETENESS 0.882450
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.796 0.965 0.813 0.910 0.986 0.770 0.981 0.457 0.961 0.953 0.434 0.998 0.0 0.0 0.0 0.0
RESIDS********** 1.66 0.68 -1.59 -1.09 -0.43 -1.77 0.50 2.88 0.71 0.78 -2.95 0.15********** 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 60
INTERCEPT(EST.OF R)= 0.030658 SLOPE (REGIST.FAC) = 1.115998 COMPLETENESS 0.896059
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********** 1.50 0.53 -1.71 -1.19 -0.50 -1.81 0.49 2.93 0.83 0.98 -2.63 0.59********** 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 15- TO 65
INTERCEPT(EST.OF R)= 0.029790 SLOPE (REGIST.FAC) = 1.148196 COMPLETENESS 0.870931
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.968 0.990 1.000 0.952 0.980 0.750 0.975 0.976 0.797 1.000 0.0 0.0 0.0 0.0
RESIDS**********512.27 -1.11 -0.63 0.01 -1.37 0.87 3.20 0.98 0.97 -2.87 0.13 11.93********** 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 15- TO 65

INTERCEPT(EST.OF R)= 0.027033 SLOPE (REGIST.FAC) = 1.314239 COMPLETENESS 0.760896
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********** 0.28 0.55 0.92 -0.77 1.09 2.89 0.02 -0.82 -5.79 -3.98 5.62********** 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.029841 SLOPE (REGIST.FAC) = 1.154177 COMPLETENESS 0.866418
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.955 0.999 0.820 0.956 0.371 0.948 0.953 0.346 0.998 0.0 0.0 0.0 0.0
RESIDS**********268.83 -0.74 -0.11 -1.49 0.72 3.04 0.79 0.75 -3.12 -0.17********** 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.030453 SLOPE (REGIST.FAC) = 1.123926 COMPLETENESS 0.889738
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********** -1.06 -0.39 -1.71 0.57 2.99 0.86 0.97 -2.70 0.47********** 0.0

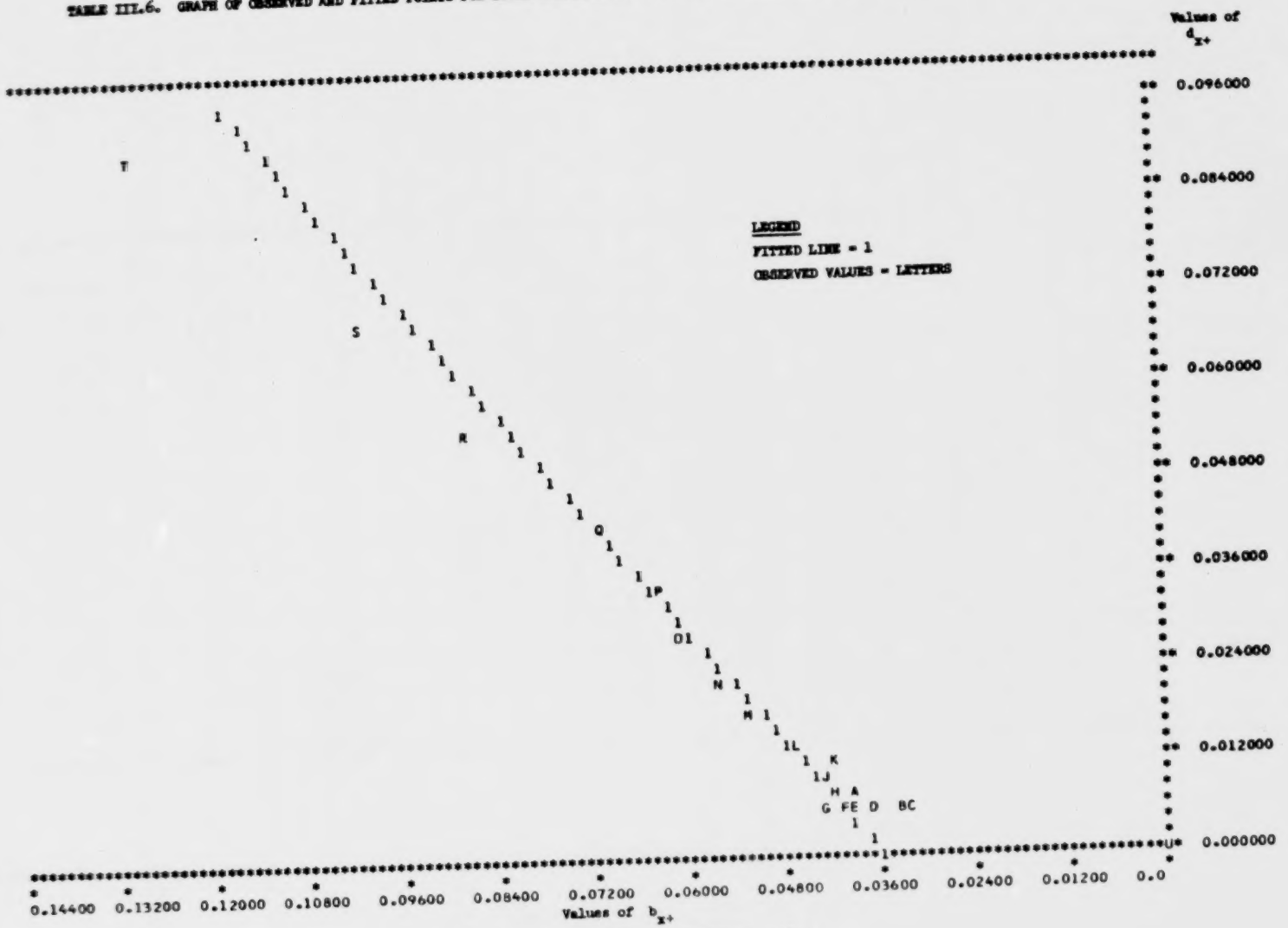
| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

TABLE III.5. RESULTS OF BRASS GROWTH-BALANCE TECHNIQUE FOR EL SALVADOR, FEMALES, 1971.

| AGES | POPULATION | DEATHS | CUMULATED POPULATION | CUMULATED DEATHS | $d_{x+}$ | $b_{x+}$ | FIT * | RESIDUALS | UNADJUSTED RATES | ADJUSTED RATES | AGES |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 64004. | 3977. | 1791458. | 14582. | 0.0081397 | 0.0390372 | 0.0426119 | -0.0035747 | 0.0621355 | 0.0527816 | 0 |
| 1 | 52144. | 1209. | 1727454. | 10605. | 0.0061391 | 0.0336183 | 0.0409125 | -0.0072942 | 0.0231850 | 0.0196947 | 1 |
| 2 | 57575. | 570. | 1675310. | 9396. | 0.0056086 | 0.0327458 | 0.0404618 | -0.0077160 | 0.0099052 | 0.0084140 | 2 |
| 3 | 61825. | 392. | 1617735. | 8826. | 0.0054557 | 0.0369032 | 0.0403319 | -0.0034287 | 0.0063447 | 0.0053896 | 3 |
| 4 | 61095. | 269. | 1555911. | 8434. | 0.0054204 | 0.0395008 | 0.0403019 | -0.0008011 | 0.0043956 | 0.0037339 | 4 |
| 5 | 285244. | 727. | 1494816. | 8165. | 0.0054622 | 0.0404199 | 0.0403375 | 0.0000825 | 0.0025494 | 0.0021656 | 5 |
| 10 | 230078. | 224. | 1209572. | 7438. | 0.0061492 | 0.0426037 | 0.0409210 | 0.0016828 | 0.0009749 | 0.0008281 | 10 |
| 15 | 184266. | 240. | 979494. | 7214. | 0.0073646 | 0.0423019 | 0.0419534 | 0.0003485 | 0.0013046 | 0.0011082 | 15 |
| 20 | 152908. | 251. | 795228. | 6973. | 0.0087688 | 0.0423997 | 0.0431462 | -0.0007466 | 0.0016444 | 0.0013969 | 20 |
| 25 | 120746. | 224. | 642320. | 6722. | 0.0104648 | 0.0426040 | 0.0445869 | -0.0019829 | 0.0018575 | 0.0015779 | 25 |
| 30 | 100635. | 272. | 521574. | 6497. | 0.0124574 | 0.0424449 | 0.0462795 | -0.0038346 | 0.0026985 | 0.0022923 | 30 |
| 35 | 95426. | 364. | 420936. | 6226. | 0.0147904 | 0.0465773 | 0.0482614 | -0.0016841 | 0.0038155 | 0.0032411 | 35 |
| 40 | 76664. | 375. | 325512. | 5862. | 0.0180078 | 0.0528676 | 0.0509944 | 0.0018732 | 0.0048936 | 0.0041569 | 40 |
| 45 | 62776. | 407. | 248848. | 5487. | 0.0220480 | 0.0560343 | 0.0544264 | 0.0016079 | 0.0064889 | 0.0055121 | 45 |
| 50 | 50563. | 453. | 186072. | 5079. | 0.0272973 | 0.0609113 | 0.0588854 | 0.0020259 | 0.0089513 | 0.0076038 | 50 |
| 55 | 36148. | 448. | 135509. | 4627. | 0.0341428 | 0.0639890 | 0.0647004 | -0.0007114 | 0.0123820 | 0.0105180 | 55 |
| 60 | 34100. | 626. | 99361. | 4179. | 0.0420594 | 0.0706996 | 0.0714252 | -0.0007256 | 0.0183459 | 0.0155841 | 60 |
| 65 | 23129. | 683. | 65261. | 3553. | 0.0544503 | 0.0876935 | 0.0819508 | 0.0057427 | 0.0295272 | 0.0250821 | 65 |
| 70 | 13473. | 814. | 42132. | 2871. | 0.0681322 | 0.1011156 | 0.0935730 | 0.0075426 | 0.0417857 | 0.0354952 | 70 |
| 75 | 9790. | 585. | 22659. | 2057. | 0.0907741 | 0.1291465 | 0.1128063 | 0.0163401 | 0.0597903 | 0.0507893 | 75 |
| UNKNOWN | ( 78.) | ( 84.) | | | | | | | | | |
| ALL | 1791458. | 14582. | | | | | | | 0.0081397 | 0.0069144 | ALL |

* Fit based on biweight regression for age-range 5+ to 60+ giving r = 0.03570 and f = 0.84946.

TABLE III.6. GRAPH OF OBSERVED AND FITTED POINTS FOR BRASS GROWTH-BALANCE METHOD, EL SALVADOR, FEMALES, 1971.




BIWEIGHT REGRESSION ESTIMATES BASED ON AGES 5+ TO 60+

(R = 0.05570 FACTOR = 0.84946)

TABLE III.7. RESULTS OF BIWEIGHT AND LEAST SQUARES REGRESSIONS THROUGH VARIOUS AGE-GROUPS, BRASS GROWTH-BALANCE TECHNIQUE. EL SALVADOR, FEMALES, 1971 (SHOWING WEIGHTS USED IN REGRESSIONS AND RESIDUALS X 1000).

| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

BIWEIGHT REGRESSION ON AGE-GROUPS 0- TO 80
INTERCEPT(EST.OF R)= 0.031780 SLOPE (REGIST.FAC) = 1.035303 COMPLETENESS 0.965901
WEIGHT 0.989 0.836 0.813 0.998 0.963 0.927 0.840 0.931 0.980 1.000 0.959 0.998 0.950 0.983 0.994 0.919 0.828 0.998 0.988 0.906 0.0
RESIDS -1.17 -4.52 -4.84 -0.52 2.11 2.99 4.46 2.90 1.54 -0.01 -2.23 -0.51 2.44 1.43 0.87 -3.14 -4.62 -0.46 -1.20 3.39 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 0- TO 80
INTERCEPT(EST.OF R)= 0.031702 SLOPE (REGIST.FAC) = 1.036278 COMPLETENESS 0.964992
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS -1.10 -4.45 -4.77 -0.45 2.18 3.06 4.53 2.97 1.61 0.06 -2.17 -0.45 2.50 1.48 0.92 -3.09 -4.59 -0.43 -1.19 3.38 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 1- TO 80
INTERCEPT(EST.OF R)= 0.031895 SLOPE (REGIST.FAC) = 1.033315 COMPLETENESS 0.967759
WEIGHT 0.0 0.840 0.818 0.997 0.969 0.936 0.858 0.940 0.984 1.000 0.958 0.997 0.957 0.986 0.995 0.922 0.838 0.998 0.989 0.909 0.0
RESIDS****** -4.62 -4.95 -0.63 2.00 2.88 4.35 2.80 1.44 -0.10 -2.32 -0.60 2.36 1.36 0.81 -3.19 -4.66 -0.47 -1.18 3.45 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 1- TO 80
INTERCEPT(EST.OF R)= 0.031797 SLOPE (REGIST.FAC) = 1.034711 COMPLETENESS 0.966453
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -4.53 -4.85 -0.54 2.10 2.97 4.44 2.88 1.53 -0.02 -2.24 -0.52 2.44 1.42 0.87 -3.14 -4.62 -0.44 -1.18 3.42 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 1- TO 60
INTERCEPT(EST.OF R)= 0.036278 SLOPE (REGIST.FAC) = 0.896691 COMPLETENESS 1.115212
WEIGHT 0.0 0.251 0.189 0.870 0.997 0.960 0.803 0.948 0.998 0.971 0.777 0.976 0.849 0.899 0.882 0.979 0.957 0.0 0.0 0.0 0.0
RESIDS****** -6.16 -6.56 -2.27 0.36 1.24 2.81 1.42 0.26 -1.06 -3.00 -0.96 2.44 1.99 2.16 -0.90 -1.29********************* 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 1- TO 60
INTERCEPT(EST.OF R)= 0.033025 SLOPE (REGIST.FAC) = 0.942096 COMPLETENESS 1.061463
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -5.19 -5.56 -1.26 1.37 2.25 3.79 2.34 1.11 -0.28 -2.32 -0.38 2.88 2.24 2.17 -1.20 -1.95********************* 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 80
INTERCEPT(EST.OF R)= 0.032922 SLOPE (REGIST.FAC) = 1.014293 COMPLETENESS 0.985909
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.973 0.917 0.974 0.998 0.994 0.932 0.987 0.980 0.996 0.999 0.912 0.838 0.999 0.994 0.881 0.0
RESIDS********************************** 1.96 3.44 1.91 0.58 -0.93 -3.11 -1.35 1.68 0.75 0.30 -3.56 -4.88 -0.46 -0.91 4.15 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 80
INTERCEPT(EST.OF R)= 0.032855 SLOPE (REGIST.FAC) = 1.015667 COMPLETENESS 0.984574
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** 2.02 3.50 1.97 0.64 -0.88 -3.06 -1.30 1.72 0.79 0.33 -3.54 -4.87 -0.46 -0.94 4.10 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 60
INTERCEPT(EST.OF R)= 0.035698 SLOPE (REGIST.FAC) = 0.849458 COMPLETENESS 1.177221
WEIGHT 0.0 0.0 0.0 0.0 0.0 1.000 0.837 0.993 0.967 0.778 0.312 0.837 0.800 0.851 0.769 0.970 0.969 0.0 0.0 0.0 0.0
RESIDS********************************** 0.08 1.68 0.35 -0.75 -1.98 -3.83 -1.68 1.87 1.61 2.03 -0.71 -0.73********************* 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 60
INTERCEPT(EST.OF R)= 0.035304 SLOPE (REGIST.FAC) = 0.861878 COMPLETENESS 1.160257
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** 0.40 1.99 0.65 -0.47 -1.72 -3.60 -1.48 2.04 1.72 2.08 -0.75 -0.86********************* 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 15- TO 65
INTERCEPT(EST.OF R)= 0.033520 SLOPE (REGIST.FAC) = 0.958705 COMPLETENESS 1.043074
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.884 0.991 0.984 0.666 0.950 0.832 0.925 0.941 0.804 0.641 0.849 0.0 0.0 0.0
RESIDS**********************************30.93****** 1.72 0.47 -0.95 -3.02 -1.12 2.08 1.38 1.22 -2.26 -3.14 1.97********** 0.0

INTERCEPT(EST.OF R)= 0.033474 SLOPE (REGIST.FAC) = 0.954158 COMPLETENESS 1.048045
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** 1.80 0.56 -0.86 -2.92 -1.01 2.21 1.52 1.39 -2.06 -2.91 2.21********** 0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.034214 SLOPE (REGIST.FAC) = 0.895572 COMPLETENESS 1.116605
WEIGHT 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.994 0.951 0.605 0.960 0.696 0.789 0.754 0.967 0.929 0.0 0.0 0.0 0.0
RESIDS********************************** 0.33 -0.98 -2.93 -0.88 2.53 2.07 2.25 -0.80 -1.18********************* 0.0

LST.SQU. REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.034055 SLOPE (REGIST.FAC) = 0.905267 COMPLETENESS 1.104647
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** 0.41 -0.92 -2.89 -0.87 2.51 2.02 2.15 -0.97 -1.43********************* 0.0

| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

stable and we might have expected additional problems with the Brass technique as a result, although destabilization typically reduces completeness estimates.

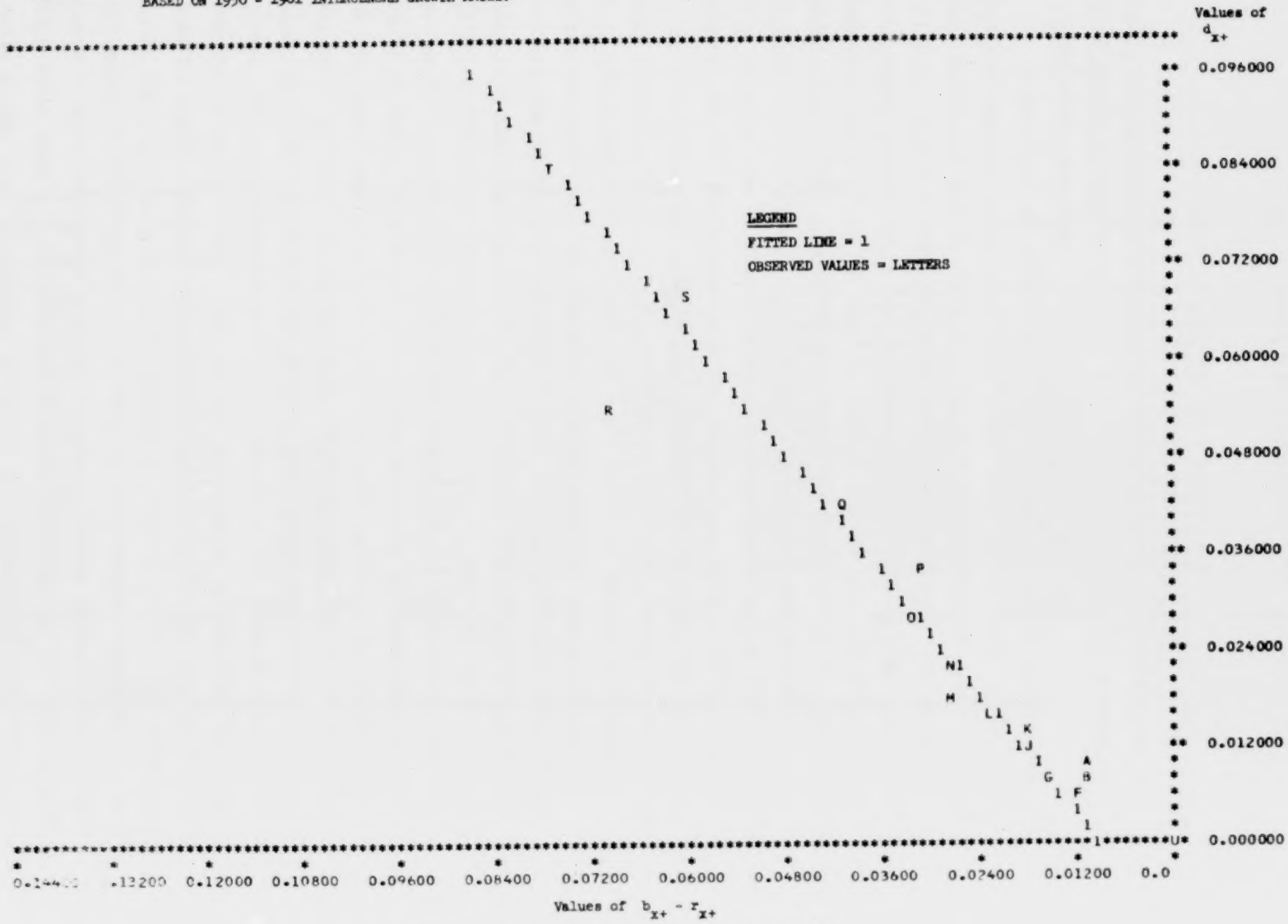
Table III.8 to III.10 give an illustration of the modified Martin approach applied to the data for females from El Salvador in 1961 using the intercensal growth rates for the period 1950 to 1961. Tables III.11 to III.13 show the technique applied to the same data, but using the intercensal growth rates for the period 1961 to 1971. Here the column labelled D(Y)/N(Y) is the same as before, namely d(y+), but the column labelled N(Y)/P(Y) is now b(y+) - r(y+) or, more correctly, the estimates thereof. The intercensal growth rates used are shown in Table III.14, from which it can be seen that for the intercensal period 1950-61 the implied growth rates at the higher ages are implausibly large. In addition these intercensal rates are fairly variable with changing age and there is quite good indirect evidence from the application of the Brass technique (and from additional evidence from the application of Preston's technique presented soon) that an assumption of stability or near stability is tenable for El Salvador in 1961. Although it is not shown here, Brass' technique also appears to work satisfactorily for 1950 albeit with a different estimated growth rate: this suggests the 1961 situation to be one of quasi-stability, at least for ages 5 or 15 and over, which would be fairly consistent with recent mortality declines as the effects of declining early childhood mortality could barely have reached this age-range by 1961. It is also clear that the position of the regression line for either the Brass technique or the Martin based modification is always heavily influenced by the few points at the highest ages for which the fit is carried out. Thus, if the intercensal growth rate at these ages is higher than in reality, the slope of the fitted line is too great, as in the examples, leading to implied overregistration. The same effect can arise in the Brass method as well as this modified approach if age at death is more overstated than age at census, as is quite likely: clearly this would make $d_{x+}$ too large and distort the slope of the line upwards, again giving rise to overhigh estimates of completeness. Of course this

TABLE III.8. RESULTS OF MODIFIED MARTIN GROWTH-BALANCE TECHNIQUE, EL SALVADOR, FEMALES, 1961 BASED ON 1950 - 1961 INTERCENSAL GROWTH RATES.

| AGES | POPULATION | DEATHS | CUMULATED POPULATION | CUMULATED DEATHS | $d_{x+}$ | $b_{x+} - r_{x+}$ | FIT * | RESIDUALS | UNADJUSTED RATES | ADJUSTED RATES | AGES |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 48114. | 4073. | 1274256. | 13654. | 0.0107153 | 0.0108925 | 0.0178438 | -0.0069512 | 0.0846530 | 0.0675772 | 0 |
| 1 | 165976. | 2837. | 1226142. | 9581. | 0.0078139 | 0.0104217 | 0.0155277 | -0.0051059 | 0.0170909 | 0.0136434 | 1 |
| 5 | 190234. | 610. | 1060166. | 6744. | 0.0063616 | 0.0114663 | 0.0143683 | -0.0029020 | 0.0032071 | 0.0025602 | 5 |
| 10 | 140538. | 214. | 869932. | 6134. | 0.0070514 | 0.0154164 | 0.0149189 | 0.0004975 | 0.0014296 | 0.0011413 | 10 |
| 15 | 125040. | 266. | 720394. | 5920. | 0.0082183 | 0.0157016 | 0.0158505 | -0.0001489 | 0.0021291 | 0.0016996 | 15 |
| 20 | 113490. | 291. | 595353. | 5654. | 0.0094972 | 0.0168475 | 0.0168714 | -0.0000239 | 0.0025679 | 0.0020499 | 20 |
| 25 | 91663. | 271. | 481863. | 5363. | 0.0111292 | 0.0180228 | 0.0181742 | -0.0001514 | 0.0029594 | 0.0023624 | 25 |
| 30 | 77711. | 315. | 390200. | 5092. | 0.0130485 | 0.0176477 | 0.0197063 | -0.0020586 | 0.0040487 | 0.0032320 | 30 |
| 35 | 72936. | 349. | 312489. | 4777. | 0.0152866 | 0.0228774 | 0.0214930 | 0.0013844 | 0.0047838 | 0.0038188 | 35 |
| 40 | 56942. | 338. | 239552. | 4428. | 0.0184843 | 0.0276580 | 0.0240457 | 0.0036123 | 0.0059327 | 0.0047360 | 40 |
| 45 | 46205. | 357. | 182610. | 4090. | 0.0223982 | 0.0281276 | 0.0271701 | 0.0009576 | 0.0077261 | 0.0061676 | 45 |
| 50 | 38616. | 385. | 136406. | 3733. | 0.0273681 | 0.0328843 | 0.0311375 | 0.0017468 | 0.0099755 | 0.0079633 | 50 |
| 55 | 26154. | 387. | 97790. | 3348. | 0.0342363 | 0.0311530 | 0.0366202 | -0.0054672 | 0.0148062 | 0.0118195 | 55 |
| 60 | 20273. | 647. | 71636. | 2961. | 0.0413300 | 0.0407236 | 0.0422830 | -0.0015594 | 0.0221160 | 0.0176549 | 60 |
| 65 | 14964. | 449. | 42363. | 2313. | 0.0546071 | 0.0693304 | 0.0528819 | 0.0164485 | 0.0299881 | 0.0239390 | 65 |
| 70 | 11205. | 504. | 27399. | 1865. | 0.0680530 | 0.0603076 | 0.0636156 | -0.0033079 | 0.0449972 | 0.0359206 | 70 |
| 75 | 6902. | 396. | 16193. | 1360. | 0.0840069 | 0.0768330 | 0.0763513 | 0.0004816 | 0.0574157 | 0.0458340 | 75 |
| UNKNOWN | ( 268.) | ( 114.) | | | | | | | | | |
| ALL | 1274256. | 13654. | | | | | | | 0.0107153 | 0.0085538 | ALL |

* Fit based on biweight regression for age-range 5+ to 60+ giving $r = 0.00929$ and $f = 0.79828$.

TABLE III.9. GRAPH OF OBSERVED AND FITTED POINTS FOR MODIFIED MARTIN GROWTH-BALANCE METHOD, EL SALVADOR, FEMALES, 1961, BASED ON 1950 - 1961 INTERCENSAL GROWTH RATES.




BIWEIGHT REGRESSION ESTIMATES BASED ON AGES 5+ TO 60+

(R = 0.01052 FACTOR = 0.74539)

TABLE III.10. RESULTS OF BIWEIGHT AND LEAST SQUARES REGRESSIONS THROUGH VARIOUS AGE-GROUPS, MODIFIED MARTIN GROWTH-BALANCE TECHNIQUE, EL SALVADOR, FEMALES, 1961, BASED ON 1950 - 1961 INTERCENSAL GROWTH RATES. (SHOWING WEIGHTS USED IN REGRESSIONS AND RESIDUALS X 1000).

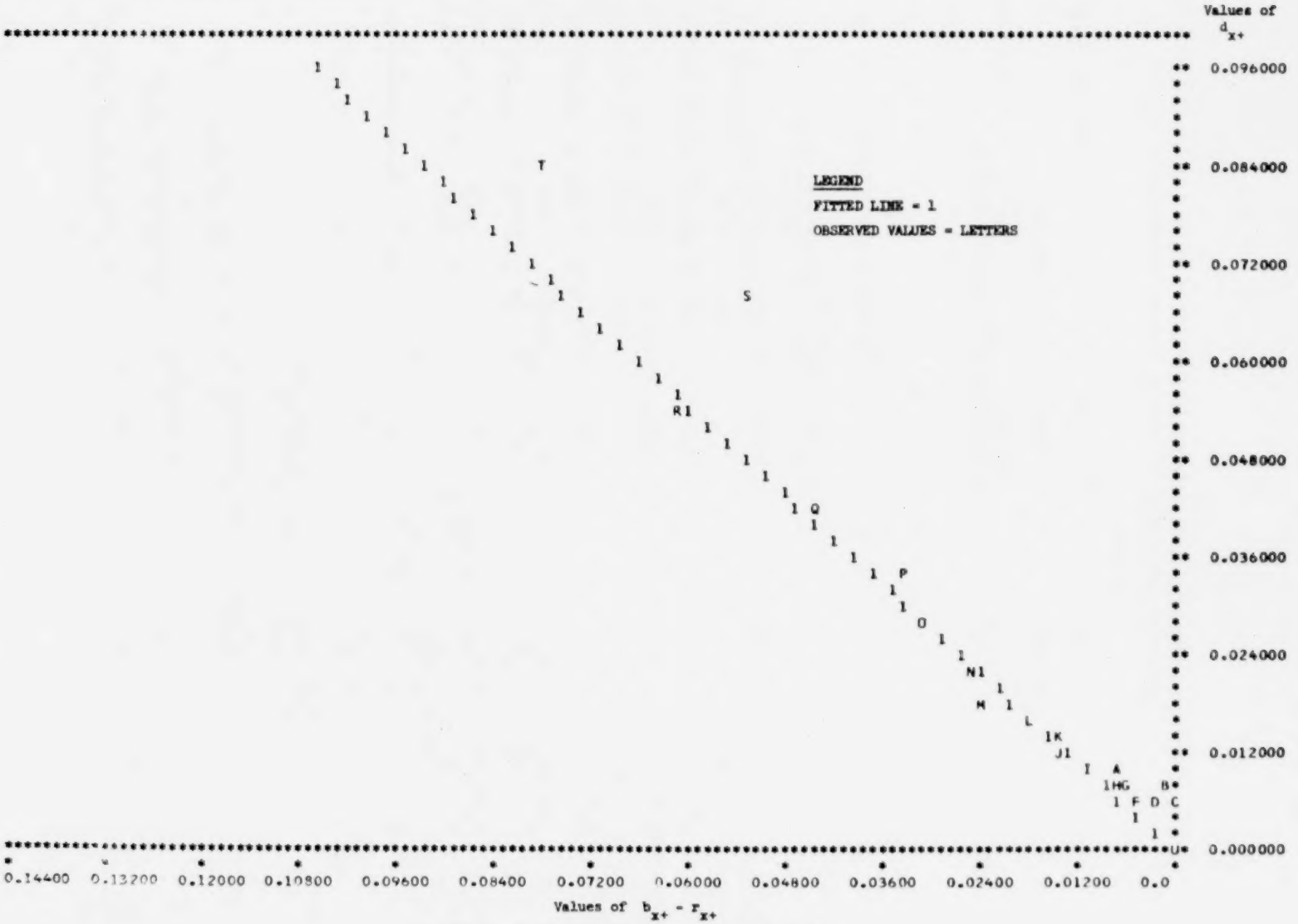
| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BIWEIGHT REGRESSION ON AGE-GROUPS 0- TO 80 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.008077 SLOPE (REGIST.FAC) = 0.805921 COMPLETENESS 1.240817 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 0.847 | 0.928 | 0.0 | 0.0 | 0.0 | 0.986 | 0.987 | 0.995 | 0.994 | 0.996 | 0.996 | 0.971 | 0.899 | 0.981 | 0.965 | 0.906 | 0.998 | 0.089 | 0.968 | 0.995 | 0.0 |
| RESIDS | -5.82 | -3.95 | 0.0 | 0.0 | 0.0 | -1.74 | 1.66 | 1.00 | 1.12 | 0.98 | -0.95 | 2.48 | 4.68 | 2.00 | 2.75 | -4.52 | -0.66 | 17.24 | -2.62 | 1.05 | 0.0 |
| LST.SQU. REGRESSION ON AGE-GROUPS 0- TO 80 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.007582 SLOPE (REGIST.FAC) = 0.858520 COMPLETENESS 1.164795 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 1.000 | 1.000 | 0.0 | 0.0 | 0.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.0 |
| RESIDS | -5.89 | -3.87 | 0.0 | 0.0 | 0.0 | -1.58 | 1.78 | 1.06 | 1.11 | 0.89 | -1.14 | 2.17 | 4.21 | 1.32 | 1.81 | -5.82 | -2.34 | 14.87 | -5.70 | -2.87 | 0.0 |
| BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 80 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.009320 SLOPE (REGIST.FAC) = 0.780060 COMPLETENESS 1.281952 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.916 | 0.996 | 1.000 | 1.000 | 1.000 | 0.963 | 0.971 | 0.841 | 0.981 | 0.948 | 0.760 | 0.992 | 0.0 | 0.953 | 0.958 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | -2.82 | 0.60 | -0.03 | 0.12 | 0.02 | -1.85 | 1.63 | 3.92 | 1.34 | 2.22 | -4.87 | -0.84 | 17.41 | -2.10 | 1.98 | 0.0 | |
| LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 80 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.008859 SLOPE (REGIST.FAC) = 0.836220 COMPLETENESS 1.195858 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 1.000 | 1.000 | 0.0 | 0.0 | 0.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | -2.71 | 0.66 | -0.03 | 0.05 | -0.14 | -2.12 | 1.24 | 3.34 | 0.54 | 1.14 | -6.34 | -2.70 | 14.81 | -5.46 | -2.27 | 0.0 | |
| BIWEIGHT REGRESSION ON AGE-GROUPS 5- TO 60 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.009290 SLOPE (REGIST.FAC) = 0.798284 COMPLETENESS 1.252687 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.672 | 0.989 | 0.999 | 1.000 | 0.999 | 0.827 | 0.920 | 0.520 | 0.961 | 0.874 | 0.130 | 0.899 | 0.0 | 0.0 | 0.0 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | -2.90 | 0.50 | -0.15 | -0.02 | -0.15 | -2.06 | 1.38 | 3.61 | 0.96 | 1.75 | -5.47 | -1.56 | ********** | 0.0 | | | |
| LST.SQU. REGRESSION ON AGE-GROUPS 5- TO 60 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.009778 SLOPE (REGIST.FAC) = 0.751789 COMPLETENESS 1.330161 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 1.000 | 1.000 | 0.0 | 0.0 | 0.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | -3.09 | 0.34 | -0.25 | -0.07 | -0.12 | -1.94 | 1.61 | 3.98 | 1.51 | 2.53 | -4.36 | -0.13 | ********** | 0.0 | | | |
| BIWEIGHT REGRESSION ON AGE-GROUPS 15- TO 65 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.010576 SLOPE (REGIST.FAC) = 0.727388 COMPLETENESS 1.374782 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.992 | 0.995 | 0.995 | 0.935 | 0.984 | 0.855 | 0.982 | 0.936 | 0.799 | 1.000 | 0.0 | 0.0 | 0.0 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | ******334.64 | -0.85 | -0.64 | -0.65 | -2.42 | 1.18 | 3.64 | 1.26 | 2.40 | -4.33 | 0.08 | 19.03 | ********** | 0.0 | | | |
| LST.SQU. REGRESSION ON AGE-GROUPS 15- TO 65 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.006076 SLOPE (REGIST.FAC) = 0.994244 COMPLETENESS 1.005789 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 1.000 | 1.000 | 0.0 | 0.0 | 0.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | ********** | 1.45 | 1.33 | 0.88 | -1.40 | 1.60 | 3.20 | -0.22 | -0.40 | -8.96 | -6.45 | 8.96 | ********** | 0.0 | | | |
| BIWEIGHT REGRESSION ON AGE-GROUPS 20- TO 60 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.010508 SLOPE (REGIST.FAC) = 0.745390 COMPLETENESS 1.341580 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.981 | 0.978 | 0.775 | 0.966 | 0.634 | 0.970 | 0.865 | 0.331 | 0.988 | 0.0 | 0.0 | 0.0 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | ********** | -0.74 | -0.78 | -2.59 | 0.98 | 3.37 | 0.92 | 1.98 | -4.87 | -0.59 | ********** | 0.0 | | | | | |
| LST.SQU. REGRESSION ON AGE-GROUPS 20- TO 60 | | | | | | | | | | | | | | | | | | | | | |
| INTERCEPT(EST.OF R)= 0.011118 SLOPE (REGIST.FAC) = 0.704867 COMPLETENESS 1.418708 | | | | | | | | | | | | | | | | | | | | | |
| WEIGHT | 1.000 | 1.000 | 0.0 | 0.0 | 0.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.0 |
| RESIDS | ********** | 0.0 | 0.0 | 0.0 | ********** | -0.96 | -0.94 | -2.67 | 0.98 | 3.51 | 1.22 | 2.48 | -4.10 | 0.47 | ********** | 0.0 | | | | | |
| AGES | 0 | 1 | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 |

TABLE III.11. RESULTS OF MODIFIED MARTIN GROWTH-BALANCE TECHNIQUE, EL SALVADOR, FEMALES, 1961 BASED ON 1961 - 1971 INTERCENSAL GROWTH RATES.

| AGES | POPULATION | DEATHS | CUMULATED POPULATION | CUMULATED DEATHS | $d_{x+}$ | $b_{x+} - r_{x+}$ | FIT * | RESIDUALS | UNADJUSTED RATES | ADJUSTED RATES | AGES |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 48114. | 4073. | 1274256. | 13654. | 0.0107153 | 0.0069454 | 0.0118176 | -0.0048721 | 0.0846530 | 0.0934307 | 0 |
| 1 | 39822. | 1349. | 1226142. | 9581. | 0.0078139 | 0.0015814 | 0.0086154 | -0.0070340 | 0.0338821 | 0.0373953 | 1 |
| 2 | 40183. | 719. | 1186320. | 8232. | 0.0069389 | -0.0007941 | 0.0076496 | -0.0084437 | 0.0178930 | 0.0197483 | 2 |
| 3 | 43630. | 482. | 1146136. | 7513. | 0.0065548 | 0.0021005 | 0.0072257 | -0.0051253 | 0.0110480 | 0.0121935 | 3 |
| 4 | 42340. | 286. | 1102506. | 7031. | 0.0063770 | 0.0045410 | 0.0070295 | -0.0024885 | 0.0067641 | 0.0074655 | 4 |
| 5 | 190234. | 610. | 1060166. | 6744. | 0.0063616 | 0.0049044 | 0.0070124 | -0.0021080 | 0.0032071 | 0.0035396 | 5 |
| 10 | 149538. | 214. | 869932. | 6134. | 0.0070514 | 0.0060967 | 0.0077738 | -0.0016771 | 0.0014296 | 0.0015779 | 10 |
| 15 | 125040. | 266. | 720394. | 5920. | 0.0082183 | 0.0073913 | 0.0090617 | -0.0016704 | 0.0021291 | 0.0023499 | 15 |
| 20 | 113490. | 291. | 595353. | 5654. | 0.0094972 | 0.0111180 | 0.0104732 | 0.0006447 | 0.0025679 | 0.0028342 | 20 |
| 25 | 91663. | 271. | 481863. | 5363. | 0.0111292 | 0.0138324 | 0.0122745 | 0.0015579 | 0.0029594 | 0.0032662 | 25 |
| 30 | 77711. | 315. | 390200. | 5092. | 0.0130485 | 0.0143880 | 0.0143927 | -0.0000046 | 0.0040487 | 0.0044685 | 30 |
| 35 | 72936. | 349. | 312489. | 4777. | 0.0152866 | 0.0184172 | 0.0168629 | 0.0015543 | 0.0047838 | 0.0052798 | 35 |
| 40 | 56942. | 338. | 239552. | 4428. | 0.0184843 | 0.0235543 | 0.0203922 | 0.0031621 | 0.0059327 | 0.0065479 | 40 |
| 45 | 46205. | 357. | 182610. | 4090. | 0.0223982 | 0.0255359 | 0.0247119 | 0.0008240 | 0.0077261 | 0.0085272 | 45 |
| 50 | 33616. | 385. | 136406. | 3733. | 0.0273681 | 0.0311327 | 0.0301972 | 0.0009356 | 0.0099755 | 0.0110099 | 50 |
| 55 | 26154. | 387. | 97790. | 3348. | 0.0342363 | 0.0336117 | 0.0377775 | -0.0041657 | 0.0148062 | 0.0163414 | 55 |
| 60 | 24273. | 647. | 71636. | 2961. | 0.0413300 | 0.0446561 | 0.0456067 | -0.0009506 | 0.0221160 | 0.0244092 | 60 |
| 65 | 14964. | 449. | 42363. | 2313. | 0.0546071 | 0.0612127 | 0.0602605 | 0.0009523 | 0.0299881 | 0.0330976 | 65 |
| 70 | 11205. | 504. | 27399. | 1865. | 0.0680530 | 0.0524829 | 0.0751006 | -0.0226177 | 0.0449972 | 0.0496629 | 70 |
| 75 | 6902. | 396. | 16193. | 1360. | 0.0840069 | 0.0782269 | 0.0927088 | -0.0144819 | 0.0574157 | 0.0633691 | 75 |
| UNKNOWN | ( 268.) | (114.) | | | | | | | | | |
| ALL | 1274256. | 13654. | | | | | | | 0.0107153 | 0.0118263 | ALL |

* Fit based on biweight regression for age-range 5+ to 60+ giving r = 0.00001 and f = 1.10369.

TABLE III.12. GRAPH OF OBSERVED AND FITTED POINTS FOR MODIFIED MARTIN GROWTH-BALANCE METHOD, EL SALVADOR, FEMALES, 1961, BASED ON 1961 - 1971 INTERCENSAL GROWTH RATES.




BIWEIGHT REGRESSION ESTIMATES BASED ON AGES 5+ TO 60+

( R = 0.00001 FACTOR = 1.10369)

TABLE III.15. RESULTS OF BIWEIGHT AND LEAST SQUARES REGRESSIONS THROUGH VARIOUS AGE-GROUPS, MODIFIED MARTIN GROWTH-BALANCE TECHNIQUE, EL SALVADOR, FEMALES, 1961, BASED ON 1961 - 1971 INTERCENSAL GROWTH RATES. (GROWING WEIGHTS USED IN REGRESSIONS AND RESIDUALS X 1000).

```
AGES        0     1     2     3     4     5    10    15    20    25    30    35    40    45    50    55    60    65    70    75    80

BIWEIGHT REGRESSION ON AGE-GROUPS  0- TO 80
INTERCEPT(EST.OF R)=-0.002807    SLOPE (REGIST.FAC) = 1.178162    COMPLETENESS 0.848779
WEIGHT 0.948 0.858 0.774 0.950 1.000 1.000 0.998 0.998 0.953 0.923 0.979 0.936 0.871 0.976 0.982 0.905 0.990 0.999 0.0    0.0    0.0
RESIDS -2.87 -4.82 -6.16 -2.82 -0.17  0.22  0.60  0.52  2.74  3.53  1.82  3.21  4.58  1.95  1.70 -3.92 -1.23 -0.32-24.89-17.94  0.0

LST.SQU. REGRESSION ON AGE-GROUPS  0- TO 80
INTERCEPT(EST.OF R)= 0.000192    SLOPE (REGIST.FAC) = 0.951294    COMPLETENESS 1.051200
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS -3.44 -6.04 -7.59 -4.33 -1.72 -1.34 -0.80 -0.62  1.89  3.05  1.78  3.68  5.78  4.04  4.91  0.85  5.15  9.07-12.45 -1.88  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS  1- TO 80
INTERCEPT(EST.OF R)=-0.002404    SLOPE (REGIST.FAC) = 1.171373    COMPLETENESS 0.853699
WEIGHT 0.0   0.843 0.755 0.940 0.999 1.000 0.999 1.000 0.961 0.932 0.984 0.943 0.881 0.980 0.984 0.901 0.990 1.000 0.0    0.000 0.0
RESIDS****** -5.08 -6.43 -3.88 -0.43 -0.05  0.33  0.26  2.49  3.29  1.60  3.01  4.40  1.79  1.57 -4.00 -1.26 -0.26-24.74-17.68  0.0

LST.SQU. REGRESSION ON AGE-GROUPS  1- TO 80
INTERCEPT(EST.OF R)= 0.000483    SLOPE (REGIST.FAC) = 0.946627    COMPLETENESS 1.056382
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -6.30 -7.85 -4.59 -1.98 -1.60 -1.06 -0.87  1.64  2.81  1.55  3.46  5.57  3.85  4.74  0.72  5.05  9.04-12.42 -1.78  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS  1- TO 60
INTERCEPT(EST.OF R)=-0.001846    SLOPE (REGIST.FAC) = 1.156624    COMPLETENESS 0.864585
WEIGHT 0.0   0.449 0.240 0.742 0.980 0.992 0.999 0.998 0.920 0.862 0.973 0.865 0.690 0.955 0.964 0.673 0.965 0.0    0.0    0.0    0.0
RESIDS****** -5.61 -6.97 -3.63 -0.99 -0.61 -0.21 -0.27  1.98  2.81  1.14  2.58  4.02  1.48  1.32 -4.14 -1.30********************  0.0

LST.SQU. REGRESSION ON AGE-GROUPS  1- TO 60
INTERCEPT(EST.OF R)=-0.002795    SLOPE (REGIST.FAC) = 1.184614    COMPLETENESS 0.844157
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS****** -4.88 -6.22 -2.87 -0.22  0.16  0.54  0.45  2.66  3.44  1.73  3.10  4.45  1.80  1.51 -4.15 -1.51********************  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS  5- TO 80
INTERCEPT(EST.OF R)= 0.001994    SLOPE (REGIST.FAC) = 0.983380    COMPLETENESS 1.016901
WEIGHT 0.0   0.0   0.0   0.0   0.0   0.892 0.922 0.930 1.000 0.992 0.998 0.981 0.890 0.977 0.952 0.959 0.960 0.721 0.0    0.638 0.0
RESIDS**************************************** -3.35 -2.83 -2.68 -0.22  0.89 -0.44  1.39  3.38  1.52  2.23 -2.05  7.07  5.52-16.43 -6.38  0.0

LST.SQU. REGRESSION ON AGE-GROUPS  5- TO 80
INTERCEPT(EST.OF R)= 0.003372    SLOPE (REGIST.FAC) = 0.892907    COMPLETENESS 1.119938
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** -4.15 -3.57 -3.32 -0.73  0.52 -0.63  1.40  3.68  2.16  3.32 -0.33  4.38  9.08-11.65 -0.16  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS  5- TO 60
INTERCEPT(EST.OF R)=-0.000009    SLOPE (REGIST.FAC) = 1.103689    COMPLETENESS 0.906052
WEIGHT 0.0   0.0   0.0   0.0   0.0   0.796 0.868 0.869 0.980 0.886 1.000 0.886 0.573 0.967 0.958 0.335 0.957 0.0    0.0    0.0    0.0
RESIDS********************************** -2.11 -1.68 -1.67  0.64  1.56 -0.00  1.55  3.16  0.82  0.94 -4.17 -0.95********************  0.0

LST.SQU. REGRESSION ON AGE-GROUPS  5- TO 60
INTERCEPT(EST.OF R)= 0.000303    SLOPE (REGIST.FAC) = 1.077381    COMPLETENESS 0.928177
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS********************************** -2.25 -1.80 -1.77  0.58  1.54  0.03  1.84  3.34  1.10  1.34 -3.58 -0.18********************  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 15- TO 65
INTERCEPT(EST.OF R)= 0.000845    SLOPE (REGIST.FAC) = 1.097186    COMPLETENESS 0.911423
WEIGHT 0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.812 0.998 0.957 0.958 0.955 0.623 0.999 0.995 0.032 0.839 0.985 0.0    0.0    0.0
RESIDS************************************************** -2.47 -0.15  0.78 -0.77  0.80  2.43  0.12  0.26 -4.80 -1.54  0.45************  0.0

LST.SQU. REGRESSION ON AGE-GROUPS 15- TO 65
INTERCEPT(EST.OF R)= 0.000882    SLOPE (REGIST.FAC) = 1.076451    COMPLETENESS 0.928979
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS************************************************** -2.34  0.01  0.97 -0.54  1.08  2.77  0.54  0.79 -4.12 -0.72  1.55************  0.0

BIWEIGHT REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.001989    SLOPE (REGIST.FAC) = 1.046598    COMPLETENESS 0.955477
WEIGHT 0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.938 0.996 0.855 0.983 0.586 0.999 0.976 0.025 0.967 0.0    0.0    0.0    0.0
RESIDS************************************************************ -0.81  0.20 -1.26  0.43  2.22  0.10  0.50 -4.21 -0.59********************  0.0

LST.SQU. REGRESSION ON AGE-GROUPS 20- TO 60
INTERCEPT(EST.OF R)= 0.002803    SLOPE (REGIST.FAC) = 0.990892    COMPLETENESS 1.009192
WEIGHT 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.0
RESIDS************************************************************ -1.10  0.00 -1.34  0.47  2.44  0.54  1.21 -3.12  0.90********************  0.0

AGES        0     1     2     3     4     5    10    15    20    25    30    35    40    45    50    55    60    65    70    75    80
```

Table III.14. Results of Preston death registration technique based on intercensal growth rates, El Salvador, Females, 1961.

| Age-range | Intercensal growth rates | | | Completeness estimates based on | | |
|---|---|---|---|---|---|---|
| | 50-61 | 61-71 | Average | 50-61 | 61-71 | Average |
| 0+ | .0279 | .0341 | .0310 | .797 | 1.063 | .930 |
| 1+ | .0277 | .0343 | .0310 | .791 | 1.070 | .930 |
| 5+ | .0262 | .0344 | .0303 | .742 | 1.056 | .899 |
| 10+ | .0236 | .0330 | .0283 | .688 | .991 | .839 |
| 15+ | .0224 | .0307 | .0266 | .670 | .898 | .784 |
| 20+ | .0232 | .0289 | .0260 | .696 | .836 | .766 |
| 25+ | .0246 | .0287 | .0266 | .740 | .836 | .788 |
| 30+ | .0258 | .0290 | .0274 | .769 | .838 | .803 |
| 35+ | .0253 | .0298 | .0276 | .762 | .847 | .805 |
| 40+ | .0266 | .0307 | .0287 | .806 | .879 | .843 |
| 45+ | .0284 | .0309 | .0296 | .844 | .886 | .865 |
| 50+ | .0293 | .0311 | .0302 | .861 | .885 | .873 |
| 55+ | .0351 | .0326 | .0338 | .942 | .910 | .926 |
| 60+ | .0366 | .0327 | .0346 | .895 | .854 | .874 |

possibility is one of the main reasons for truncating the upper ages at 60 or thereabouts, but any spill-over effects in the two or three highest ages used will inevitably lead to distorted estimates, although this might be apparent as an obvious departure from linearity elsewhere. However, by the time fitting is restricted to a quite limited age range, such as 15 to 65 (11 points) or even 20 to 60 (9 points), with some of the points grouped fairly close to the origin at the younger ages, it really becomes quite difficult to identify such distortions visually. Visual distinction between a line of unit slope and one with a slope of 0.9 is not always clear-cut, especially with points clustered in the lower corner of the graph. In addition, fitting procedures are not very robust with so few points.

Whether for the reasons suggested above or because of other errors present in the data, it is clear from Tables III.8 to III.13 that the modified technique is of no help for 1961-based estimates especially when using the 1950-61 observed intercensal growth rates. Using the 1961-1971 observed intercensal growth rates, the estimates of completeness are at least plausible and close to Brass-technique results when based on the age-range 1 to 60 or a subset thereof. More generally, the estimates obtained from this modified procedure have not often proved very helpful, there being a very noticeable tendency for estimated completeness to exceed 100%, sometimes seriously. Of course similar problems have been experienced with the original Brass technique as well and both $d_{y+}$ and $b_{y+}$ are estimated identically for the two techniques and are thus susceptible to the same forms of error. Nevertheless the errors in the modified technique for El Salvador cannot be entirely due to the estimates of $d_{y+}$ and $b_{y+}$, as the Brass technique works well. Thus, in this case it is clearly the intercensal growth rates which are problematic. There will often be such problems in using intercensal growth rates as estimators of growth rates at the time of one of the censuses, arising both from differential enumeration levels and from fairly rapid changes in the rates over time, especially during the earlier stages of the demographic transition. In addition the differencing

of $b_{x+}$ and $r_{x+}$, which are both subject to errors, concentrates all the errors into a smaller number, causing unreliability.

The last of the recently developed techniques for assessing completeness of death registration due to Preston is now considered. Table III.14 shows the completeness estimates derived for ages above the indicated ages, using the two sets of intercensal growth rate estimates. Here, for ages y and over, the value of $r_{y+}$ is applied throughout the age-range y and over, as the estimator of the stable or quasi-stable population growth rate. Table III.15 shows the application of the Preston technique for various fixed rates of natural increase, ranging from 1% to 3.5% per annum. The estimated completeness increases monotonically with increasing values of r: this relation is obvious from the expression given earlier for C, as both r and the integral are monotically increasing with increasing r. What is most noticeable about this particular example (again El Salvador females 1961) is the astonishing consistency of the completeness estimates based on an r of .030, with estimated completeness never leaving the range 86 to 88% for starting ages between 0 and 55. Of course, we should not make too much of this consistency, as use of a value of r of .029 or .031 would also give very consistent estimates at slightly different levels of estimated completeness, although in this particular case .030 is almost certainly very close the optimal value for consistent estimates. A comparison with the estimates using the Brass technique (Table III.4) shows a close correspondence with very similar estimates of completeness and the Brass technique giving estimates of r which are very close to .030. As noted earlier the Preston technique gives high weight to deaths at the older ages. As a result it gives, at least in this instance, consistent results even using all ages zero and over, unlike the Brass technique. We have noted that the consistency is quite remarkable and must stress that this is by no means always the case. There are many instances in other countries where the estimated completeness levels are inconsistent or implausible and where the Preston and Brass techniques provide somewhat different estimates. This is to be expected as both techniques make use of similar data and assumptions, although differing in emphasis, with Preston's technique avoiding the direct estimation of $b_{y+}$ but instead placing far more reliance on the distribution of deaths.

TABLE III.15. RESULTS OF PRESTON DEATH REGISTRATION TECHNIQUE WITH ARBITRARY GROWTH RATES, EL SALVADOR, FEMALES, 1961.

| AGE RANGE | DEATH DISTRIBUTION | VALUES OF INTEGRALS | | | | | | COMPLETENESS ESTIMATES | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | r=.010 | r=.015 | r=.020 | r=.025 | r=.030 | r=.035 | r=.010 | r=.015 | r=.020 | r=.025 | r=.030 | r=.035 |
| 0+ | 0.29830 | 1.35933 | 1.64918 | 2.05627 | 2.63260 | 3.45442 | 4.63389 | 0.38503 | 0.46374 | 0.56591 | 0.69975 | 0.87666 | 1.11252 |
| 1+ | 0.09882 | 1.49577 | 1.89460 | 2.45318 | 3.24138 | 4.36117 | 5.96185 | 0.38739 | 0.46603 | 0.56775 | 0.70056 | 0.87546 | 1.10776 |
| 2+ | 0.05266 | 1.56069 | 2.00988 | 2.63679 | 3.51806 | 4.76500 | 6.53989 | 0.38906 | 0.46716 | 0.56787 | 0.69890 | 0.87083 | 1.09830 |
| 3+ | 0.03530 | 1.59781 | 2.07446 | 2.73718 | 3.66506 | 4.97245 | 6.82528 | 0.39186 | 0.46952 | 0.56935 | 0.69876 | 0.86796 | 1.09096 |
| 4+ | 0.02097 | 1.62215 | 2.11563 | 2.79905 | 3.75193 | 5.08879 | 6.97501 | 0.39675 | 0.47429 | 0.57363 | 0.70197 | 0.86914 | 1.08865 |
| 5+ | 0.04468 | 1.63196 | 2.13049 | 2.81809 | 3.77275 | 5.10626 | 6.97938 | 0.40202 | 0.47944 | 0.57829 | 0.70556 | 0.87074 | 1.08680 |
| 10+ | 0.01566 | 1.60976 | 2.07733 | 2.70891 | 3.56715 | 4.73985 | 6.35047 | 0.42996 | 0.50644 | 0.60251 | 0.72408 | 0.87904 | 1.07795 |
| 15+ | 0.01950 | 1.55132 | 1.96204 | 2.50529 | 3.22775 | 4.19344 | 5.49036 | 0.45310 | 0.52709 | 0.61855 | 0.73233 | 0.87482 | 1.05438 |
| 20+ | 0.02134 | 1.49922 | 1.86063 | 2.32882 | 2.93836 | 3.73558 | 4.78279 | 0.47413 | 0.54490 | 0.63101 | 0.73636 | 0.86602 | 1.02646 |
| 25+ | 0.01987 | 1.45060 | 1.76765 | 2.17003 | 2.68296 | 3.33956 | 4.18334 | 0.50149 | 0.56956 | 0.65108 | 0.74920 | 0.86792 | 1.01223 |
| 30+ | 0.02304 | 1.40141 | 1.67598 | 2.01745 | 2.44380 | 2.97810 | 3.65002 | 0.52378 | 0.58803 | 0.66381 | 0.75358 | 0.86037 | 0.98796 |
| 35+ | 0.02555 | 1.35663 | 1.59385 | 1.88305 | 2.23682 | 2.67100 | 3.20550 | 0.54516 | 0.60519 | 0.67494 | 0.75627 | 0.85146 | 0.96327 |
| 40+ | 0.02474 | 1.31530 | 1.51930 | 1.76316 | 2.05551 | 2.40699 | 2.83072 | 0.58281 | 0.63993 | 0.70532 | 0.78042 | 0.86691 | 0.96685 |
| 45+ | 0.02614 | 1.27393 | 1.44639 | 1.64857 | 1.88622 | 2.16620 | 2.49686 | 0.61356 | 0.66655 | 0.72634 | 0.79399 | 0.87070 | 0.95791 |
| 50+ | 0.02821 | 1.23442 | 1.37809 | 1.54337 | 1.73392 | 1.95404 | 2.20882 | 0.64155 | 0.68983 | 0.74355 | 0.80344 | 0.87035 | 0.94523 |
| 55+ | 0.02836 | 1.19710 | 1.31479 | 1.44773 | 1.59815 | 1.76863 | 1.96214 | 0.67479 | 0.71848 | 0.76644 | 0.81914 | 0.87717 | 0.94114 |
| 60+ | 0.04742 | 1.16009 | 1.25335 | 1.35688 | 1.47196 | 1.60003 | 1.74272 | 0.66164 | 0.69806 | 0.73750 | 0.78025 | 0.82664 | 0.87705 |

5. Discussion

This section on growth-balance and related methods for assessing death registration completeness is concluded first by drawing attention to the fact that El Salvador was deliberately chosen as a case where three successive data sets were available, and that the examples are taken from the central data set, which allows the maximum number of estimates to be obtained. For many countries only a single data set is available, severely reducing the variety of techniques which can be tried: in such cases only the Brass growth-balance method and the Preston technique with arbitraty growth rates can be used. Unless an external estimate of an appropriate value of r is available for the Preston technique it becomes difficult to get estimates therefrom, the main hope being through a search for a value of r that maximises consistency of the estimates by age. As yet we do not have enough experience with this approach to choosing an appropriate r value; for El Salvador 1961 it clearly works but elsewhere the results are often less clear. It is also not yet clear whether a search for maximal consistency virtually forces the Preston technique to give the same results as Brass', although experience so far suggests this may be true.

The second and final set of concluding remarks on these techniques take the form of a wider review of the estimates for El Salvador, albeit in limited summary form. This is an important task. The findings for females in 1961 would have much more weight if they were consistent with those for 1950 and 1971 and also with the results for males (and both sexes combined to a lesser extent for consistency). Table III.16 shows the results of applying the various registration completeness techniques. This kind of consistency analysis is extremely important and similar tables will be one of the important items of information to be presented in the detailed country analyses. These results present a fairly coherent picture, with the between sex differences being generally quite small. It would appear that death registration for adults (or perhaps those aged 5 and over) was about 80% complete in 1950; perhaps 88% in 1961 (ignoring results

Table III.16. Review of Death Registration Completeness Estimates

**EL SALVADOR**

| Year | Technique | Indicator | Both Sexes | Males | Females |
|---|---|---|---|---|---|
| 1950 | - | RAW $\overset{\circ}{e}_o$ | 50.32 | 49.06 | 51.60 |
| | BRASS | COMP* | 71/78 | 69/72 | 72/86 |
| | | r | .022/.025 | .021/.023 | .023/.028 |
| | PRESTON-ARB* | COMP | 82/85 (r=.025) | 79/84 (r=.025) | 81/86 (r=.025) |
| | " 50-61 | COMP | 73/85 | 72/83 | 75/88 |
| | MARTIN-50-61 | COMP | 95/97 | 87/109 | 95/108 |
| | | r | .004/.005 | .003/.007 | .003/.007 |
| 1961 | - | RAW $\overset{\circ}{e}_o$ | 57.81 | 55.81 | 59.83 |
| | BRASS | COMP | 81/88 | 76/87 | 87/88 |
| | | r | .027/.030 | .024/.029 | .030 |
| | PRESTON-ARB | COMP | 87/90 (r=.030) | 88/92 (r=.030) | 86/88 (r=.030) |
| | " 50-61 | COMP | 68/75 | 69/74 | 67/74 |
| | " 61-71 | COMP | 87/111 | 89/117 | 84/105 |
| | MARTIN 50-61 | COMP | 122/133 | 112/125 | 125/137 |
| | " " | r | .009/.010 | .008/.010 | .009/.011 |
| | " 61-71 | COMP | 83/87 | 72/82 | 86/96 |
| | " " | r | -.003/-.001 | -.005/-.004 | -.001/+.001 |
| 1971 | - | RAW $\overset{\circ}{e}_o$ | 61.95 | 59.42 | 64.57 |
| | BRASS | COMP | 104/123 | 104/128 | 104/118 |
| | | r | .032/.036 | .031/.036 | .033/.036 |
| | PRESTON-ARB | COMP | 101/106 (r=.033) | 114-119 (r=.035)<br>105/111 (r=.033) | 107/109 (r=.035)<br>98/101 (r=.033) |
| | 61-71 | COMP | 94/112 | 98/120 | 88/104 |
| | MARTIN 61-71 | COMP | 109/131 | 108/124 | 110/138 |
| | " " | r | .001/.005 | .000/.003 | .003/.007 |

* ARB = arbitrary growth rates

* COMP= completeness estimate

based on 1950/61 intercensal growth rates) and probably fairly complete by 1971, although this last with reservations. We give the resulting values for $\overset{\circ}{e}_5$ and $\overset{\circ}{e}_o$ (in which we place less trust) below for both sexes combined. These results look plausible, and are probably the best obtainable for El Salvador. Nevertheless, caution is needed because it is

| Year | Estimated completeness | Correction factor | Uncorrected $\overset{\circ}{e}_o$ | Uncorrected $\overset{\circ}{e}_5$ | Corrected $\overset{\circ}{e}_o$ | Corrected $\overset{\circ}{e}_5$ |
|---|---|---|---|---|---|---|
| 1950 | 80% | 1.25 | 50.32 | 57.13 | 44.6 | 53.0 |
| 1961 | 88% | 1.136 | 57.81 | 62.77 | 54.7 | 60.6 |
| 1971 | 96% | 1.04 | 61.95 | 64.22 | 61.1 | 63.6 |

conceivable that the apparent improvement in registration over time is solely due to increasing departures over time from the assumptions of the techniques used. Undoubtedly the findings for 1971 are not strongly based. It is possible that registration could have been around 80% completeness in El Salvador throughout, although the estimates of improving registration appear more plausible.

6. Validation against external models

a) Infant and childhood mortality

The validation of infant and childhood mortality data presents the most difficult part of the data assessment. The growth-balance and related methods at best only provide estimates of completeness of adult registration (5 or 10 and over being the maximal age range and 15 or 20 and over being more usual). There is a lot of evidence that infant and early childhood mortality is highly variable for given levels of adult mortality. In addition there is some evidence, and a lot of opinion, that childhood deaths are differentially underregistered. As an example we show figures from EDENH (Ortega and Rincón (1975)) comparing rates based on registered deaths by age from vital statistics with those from EDENH (see Table III.17).

Table III.17. Comparison of registration based rates with rates from prospective survey, Honduras.

| Age | Ratio of registered rate to EDENH rate | Age (years) | Ratio of registered rate to EDENH rate |
|---|---|---|---|
| < 1 day | .025 | 5-9 | .753 |
| 1-6 days | .226 | 10-19 | .702 |
| 7-29 days | .263 | 20-29 | .824 |
| 30-181 days | .509 | 30-39 | .864 |
| 182 days - 1 year | .437 | 40-49 | 1.048 |
| 1st year | .312 | 50-59 | .771 |
| 1 year | .405 | 60-69 | 1.097 |
| 2 years | .459 | 70 + | 1.295 |
| 3 years | .570 | 5 and over | .869 |
| 4 years | 1.140 | | |
| 1 to 4 years | .497 | | |

NOTE: for registration data the groups in the first year of life were: <1 day; 1-6 days; 7-27 days; 28 days to 5 months; 6 to 11 months.

It is clear from this table that deaths below age 5 were considerably less well registered than at ages 5 and over, especially in the first month of life. Table III.17 also suggests that assumptions of equal underregistration by age above age 5 are dubious, especially owing to overstatement of age of decedents. Thus it suggests caution with the growth-balance and related adult mortality correction methods, which all necessarily assume constant underregistration of deaths with age. Of course the figures presented in Table III.17 are not ratios of deaths but of death rates, and thus to an extent, confound population age distribution errors with underregistration of deaths. Nevertheless the figures presented from EDENH are one of the better items of data

on differential underregistration by age available for a developing country.

This likelihood of differential underregistration of childhood deaths makes it highly desirable to have a means of assessing such underregistration. One approach described later is to compare registration estimates with those from other sources, most notably from questions on childhood survival. Another is through fitting model life-tables to non-childhood ages and comparing reported childhood mortality with that of the fitted model, but this is difficult given the aims of the life table project, although we do describe such a technique later. In addition, a fairly exhaustive search for suitable models of infant and childhood mortality, to use for suggesting levels of completeness of early childhood mortality or infancy registration, was unproductive. Most of the models were clearly inappropriate. Bourgeois-Pichat's (1946 and subsequent) biometric approach is solely based on the first year of life and thus inappropriate. The mathematical curves suggested by Opperman (1870), Steffensen (1930) and Harper (1936) are not appropriate, as they have 3 or 4 parameters, which is excessive in view of the small number of observations typically available and the need to add parameters for corrective or at least diagnostic purposes. Extensive experimentation was also carried out with transformations in the hope of obtaining a simple model: some of the attempts were promising but did not appear sufficiently so to use at this stage.

This left one other possibility, which was to use the model of infant and childhood mortality developed by Hogan (1976). This model was based almost entirely on experience of developed countries with infant mortality of less than 100 per thousand. Hogan's model is essentially a principal components one, but is fitted using robust estimation procedures due to Tukey. The model was fitted to observed ${}_nM_x$ values for age-groups 0, 1-4, 5-9 and 10-14, and took the form

$b_x = CA_x + RM.CM_x + Rl.Cl_x$, where $b_x = /\overline{100{,}000\ {}_nM_x}$, CA, CM and Cl are the 'principal components' and RM and Rl are the fitter parameters which are country or case specific. The values of $CA_x$, $CM_x$ and $Cl_x$ are given below.

Table III.17a. Values of CA, CM and C1 for Hogan's model of early mortality.

| Age group | Males | | | Females | | |
|---|---|---|---|---|---|---|
| | CA | CM | C1 | CA | CM | C1 |
| 0 | 72.8 | -404.4 | 27.9 | 64.8 | -381.5 | 29.9 |
| 1-4 | 17.8 | -153.6 | -66.0 | 16.9 | -158.0 | -59.9 |
| 5-9 | 10.6 | - 65.9 | -29.9 | 9.5 | - 73.5 | -25.0 |
| 10-14 | 9.4 | -47.8 | -13.1 | 8.3 | - 58.0 | -19.0 |

For applications to both sexes combined, the averages of the male and female values were used. After a few preliminary experiments it became clear that use of ordinary least-squares fitting to solve for parameters was not useful, with the parameters being highly sensitive to errors. In line with the original fitting procedures used by Hogan a sequential fit was adopted. First the first component was fitted on its own by obtaining four point estimates of RM (one from each age group) and then choosing a central value from these using a Tukey biweight centring routine due to McNeil (1977). Using this estimate point estimates of R1 were calculated and found. Such a fitting procedure is moderately robust but almost certainly inefficient. The main reason for its adoption was that it was in line with the way Hogan's original model was estimated. The main objection to both these procedures and Hogan's model proves to be related to the fit being carried out to a transformation of the observations. The square-root transform was adopted by Hogan as a means of reducing heteroscedasticity. It does achieve this, but when transformed back to the original space, quite small errors in the estimate of the square root of large values lead to large errors in the estimates of the actual values. Whilst these problems could be partially overcome by an approrpiate approximate weighted least

squares fitting procedure, this seemed inappropriate unless the model were itself re-estimated taking account of this problem, which could not be done. These fitting problems and perhaps the limited experience upon which Hogan's model was based rendered the technique unpromising for the model life table project.

We give an example of the fitting procedure as applied to El Salvador Females 1961 at Table III.18. This illustrates some of the problems experienced with this approach to evaluation of patterns of childhood mortality. The fits to the data are not close and in general it proves impossible to draw inferences from the results.

b) Model life tables

While it is clearly unacceptable to use model life-tables, or estimates heavily dependent upon them, for inputs to a project to produce new model life-tables, it is still very useful to compare observed or apparent patterns of mortality with existing model life-tables as a screening procedure. In view of the results of the methodological work described in a later section the models chosen for these comparisons were the ones based on the first two principal components of the set of 326 life-tables originally used by Coale and Demeny. The values for and description of these principal components are given at Table IV.10 in Section IV. At Table III.19 are shown examples of fits of principal components based model life-tables to the data for El Salvador females for 1961. In addition to fitting over the entire age-range 0 to 80, a fit was carried out over the age-range 5-59, in an attempt to overcome any possible distortions due to larger errors at the extreme ages. For consistency with the estimation procedures used, these fits have been obtained by ordinary least-squares methods. For El Salvador the results do give some insights into the data, with the fitted values for age-groups 1 to 4 and 5 to 9 being far too low relative to the model and those for the higher ages being too high. To ensure that this is a real effect we have also tried fitting the Coale Demeny models to these data using $\overset{\circ}{e}_{10}$ to choose the model. Whichever model life-table used, the indications are always similar: mortality in the age range 1 to 9 and especially

Table III.18. Hogan analysis of infant and childhood mortality, El Salvador, Females, 1971.

| Age | Point fit of | | | Fitted values | |
|---|---|---|---|---|---|
| | $_nm_x$ | RM | RI given RM = -0.0997 | RM = -0.0997 RI = 0.0 | RM = -0.0997 RI = -0.1037 |
| 0 | .08465 | -0.0713 | -0.3618 | .10573 | .09945 |
| 1-4 | .01709 | -0.1547 | -0.1451 | .01066 | .01510 |
| 5-9 | .00321 | -0.1144 | -0.0433 | .00283 | .00377 |
| 10-14 | .00143 | -0.0630 | 0.1118 | .00198 | .00258 |
| Centre | | -0.0997 | -0.1037 | | |

TABLE III.19. PRINCIPAL COMPONENTS MODEL LIFE-TABLE FIT (SEE SECTION IV) TO EL SALVADOR, FEMALES, 1961, SHOWING VALUES OF PARAMETERS A AND B FOR EACH OF TWO FITS, ONE TO THE ENTIRE AGE-RANGE AND THE OTHER TO AGE-RANGE 5-59.

| | | FIT TO ALL AGES | | FIT THROUGH 5-59 | |
|---|---|---|---|---|---|
| | | A= -0.71114 | | A= -0.49450 | |
| | OBSERVED | B= -0.54552 | | B= 0.11921 | |
| AGE | Q-VALUES | FIT | RESIDUALS | FIT | RESIDUALS |
| 0 | 0.07984 | 0.07090 | 0.00894 | 0.07104 | 0.00880 |
| 1 | 0.06715 | 0.03500 | 0.03216 | 0.02928 | 0.03788 |
| 5 | 0.01591 | 0.01178 | 0.00413 | 0.01011 | 0.00580 |
| 10 | 0.00707 | 0.00882 | -0.00174 | 0.00742 | -0.00034 |
| 15 | 0.01062 | 0.01330 | -0.00268 | 0.01128 | -0.00066 |
| 20 | 0.01278 | 0.01800 | -0.00522 | 0.01515 | -0.00237 |
| 25 | 0.01475 | 0.02037 | -0.00562 | 0.01770 | -0.00296 |
| 30 | 0.02008 | 0.02251 | -0.00243 | 0.02038 | -0.00029 |
| 35 | 0.02369 | 0.02562 | -0.00193 | 0.02418 | -0.00049 |
| 40 | 0.02937 | 0.02992 | -0.00056 | 0.02946 | -0.00009 |
| 45 | 0.03805 | 0.03617 | 0.00188 | 0.03767 | 0.00038 |
| 50 | 0.04902 | 0.04829 | 0.00073 | 0.05225 | -0.00324 |
| 55 | 0.07178 | 0.06449 | 0.00728 | 0.07297 | -0.00119 |
| 60 | 0.10523 | 0.09774 | 0.00749 | 0.11145 | -0.00622 |
| 65 | 0.14171 | 0.14670 | -0.00499 | 0.16929 | -0.02758 |
| 70 | 0.20311 | 0.22794 | -0.02483 | 0.26061 | -0.05750 |
| 75 | 0.25215 | 0.33875 | -0.08660 | 0.38716 | -0.13501 |

1 to 4 is exceptionally high - perhaps due to overstatement of age at death for young children; mortality at the highest ages is low, perhaps due again to overstatement pushing the deaths to higher ages. Thus existing model life-tables provide a useful diagnostic tool, but are not acceptable as a corrective one.

c) Makeham curves

Actuaries often use mathematical curves to graduate life-tables (see e.g. Benjamin and Haycocks (1970)). The most commonly used curves for this purpose are Gompertz and Makeham curves, although for more recent experience some of Perk's modifications, especially the logistic, are used, perhaps in conjunction with another curve, as for The English National Life Tables. Gompertz and Makeham curves are fairly easy to handle and to use for point fitting procedures, as here. One of the main purposes in using Makeham curves is to diagnose errors in the data. For this purpose successive point-fitting through sequences of ages is at least as satisfactory as adopting a more sophisticated non-linear least squares fitting procedure and considerably less complex.

The Makeham curve is given by $\mu x = A + BC^x$

or $l_x = d.a^x \cdot b^{c^x}$, when $a = \exp(-A)$; $b = \exp(-B/\ln C)$; $d = b \cdot l_o$.

For a point fit through $l_x$, $l_{x+5}$, $l_{x+10}$ and $l_{x+15}$

we get $c^5 = \dfrac{\Delta \ln({}_5p_{x+5})}{\Delta \ln({}_5p_x)}$, where $\Delta$ denotes forward differencing,

giving $b = \exp - \left[ \frac{(\Delta \ln {}_5p_x)^3}{(\Delta^2 \ln {}_5p_x)^2} \right]$

$$a = \exp \left\{ 0.2 \left[ \frac{(\Delta \ln {}_5p_x)^2}{\Delta^2 \ln {}_5p_x} + \ln {}_5p_x \right] \right\}$$

and $d = l_x / b$.

Analogous solutions are available for a point fit through four successive ages at ten-year intervals. A series of point fits through successive ages proves to be a powerful diagnostic tool for helping to assess quality of information at the higher ages. As an example, Table III.20 shows the results of such point fits to the life-table for females for El Salvador 1961. The most sensitive indicator of differential deficiencies in the life-table values is the value of C. For the fits through four ages at five-year intervals starting at age 50 and upwards there is a clear cyclical pattern which is almost certainly indicative of age mis-statement effects. The estimates of C are unaffected by constant scaling of the death rates, at least with the life-table method used. Thus any indications of deficiencies are of relative deficiencies or errors, which do not include any overall constant underregistration of deaths effects. This makes the approach even more powerful as it can identify probable departures from one of the crucial assumptions of the various techniques for assessment of completeness of registration, namely of constancy with age. In this sense the Makeham analysis complements the other techniques very well. It should also be

clear from Table III.20 that values of C of around 1.08 are acceptable (this is generally true) and that highly deviant values (perhaps outside the range 1.05 to 1.12 or even less) lead to patently absurd values for the other parameters. As may be expected the point fits through four successive ages ten years apart are better behaved as they average out some of the digital preference effects. The last panel of Table III.20 shows the actual fitted values obtained from three of these ten-year age-group fits. The fit through ages 55, 65, 75 and 85 is patently absurd. The other two fits are more plausible, but in this particular case the one through ages 30, 40, 50 and 60 seems most satisfactory. The implications of this fit are that $l_x$ is underestimated between 65 and 75 and then progressively overestimated. This is probably most consistent with over-statement of ages, perhaps especially for decedents, but could also reflect some omission of deaths at high ages. The Makeham technique thus proves a powerful diagnostic tool. It also has the advantage of being corrective, by taking the fitted values at the higher ages as being a likely better approximation to the real experience than the observed values. Although these analyses are commenced as low as age 20 it would never be sensible to graduate these younger ages by this approach because Makeham curves fit mortality at the older ages. Another advantage of taking a Makeham fit for, say, ages 30 and over in this instance is the removal of the cyclical element in the mortality brought about by digital preference. Thus in many instances, including El Salvador Females, 1961, one of the Makeham fits will be adopted to improve life-table estimates for the higher ages.

7. Validation against other sources

For many countries there are alternative sources of mortality information over and above registration and census data, or occasional survey data on deaths and population. The most frequent information of this kind is on childhood mortality, based on reports of children ever born and children surviving using Brass' technique, or one of the later modifications of Sullivan and Trussell. Less frequently estimates are available for adult mortality, based in particular on reports of orphanhood, but also occasionally

TABLE III.20. RESULTS OF FITTING MAKEHAM CURVES TO EL SALVADOR FEMALES 1961, OBSERVED DATA.

PANEL I: POINT FITS TO FOUR SUCCESSIVE AGES FIVE YEARS APART, SHOWING VALUES OF THE PARAMETERS A, B, C AND D AND DIFFERENCES FROM THE CENTRAL VALUES OF A, B AND C, DENOTED BY DELA, DELB AND DELC.

| AGE-RANGE | A | DELA | B | DELB | C | DELC | D |
|---|---|---|---|---|---|---|---|
| 20 TO 35 | 0.997660 | 0.000650 | 0.999332 | 0.084991 | 1.222185 | 0.138804 | 83039. |
| 25 TO 40 | 0.993669 | -0.003342 | 0.948796 | 0.034454 | 0.925365 | -0.158016 | 86344. |
| 30 TO 45 | 0.997212 | 0.000201 | 0.989194 | 0.074852 | 1.096138 | 0.012758 | 81596. |
| 35 TO 50 | 0.997363 | 0.000353 | 0.980293 | 0.065952 | 1.090383 | 0.007002 | 80683. |
| 40 TO 55 | 1.000574 | 0.003563 | 0.888018 | -0.026323 | 1.049810 | -0.033570 | 86957. |
| 45 TO 60 | 0.994317 | -0.002694 | 0.990789 | 0.076448 | 1.161396 | 0.078016 | 75649. |
| 50 TO 65 | 0.999350 | 0.002340 | 0.912794 | -0.001547 | 1.086665 | 0.003285 | 78988. |
| 55 TO 70 | 1.040528 | 0.043517 | 0.131010 | -0.783331 | 1.025538 | -0.057843 | 523359. |
| 60 TO 75 | 0.988468 | -0.008543 | 0.934307 | 0.019965 | 1.122583 | 0.039202 | 68119. |
| 65 TO 80 | 0.875081 | -0.121930 | 0.028306 | -0.886035 | 0.969314 | -0.114066 | 2011815. |
| 70 TO 85 | 0.968872 | -0.028139 | 0.927928 | 0.013586 | 1.139531 | 0.056151 | 52673. |
| CENTRES | 0.997011 | | 0.914341 | | 1.083381 | | |

PANEL II: AS PANEL I BUT FITS TO FOUR SUCCESSIVE AGES TEN YEARS APART.

| AGE-RANGE | A | DELA | B | DELB | C | DELC | D |
|---|---|---|---|---|---|---|---|
| 20 TO 50 | 1.000737 | 0.000544 | 0.928280 | 0.005788 | 1.039381 | -0.032934 | 89395. |
| 25 TO 55 | 0.998574 | -0.001619 | 0.976693 | 0.054401 | 1.065888 | -0.006427 | 83860. |
| 30 TO 60 | 0.997440 | -0.002753 | 0.985838 | 0.063346 | 1.087158 | 0.014843 | 81874. |
| 35 TO 65 | 0.996656 | -0.003536 | 0.988399 | 0.065907 | 1.105998 | 0.033683 | 80022. |
| 40 TO 70 | 0.996936 | -0.003257 | 0.974718 | 0.052226 | 1.095107 | 0.022792 | 79223. |
| 45 TO 75 | 1.000659 | 0.000466 | 0.909655 | -0.012836 | 1.072313 | -0.000002 | 82396. |
| 50 TO 80 | 1.004498 | 0.004305 | 0.813344 | -0.109147 | 1.061772 | -0.010543 | 88646. |
| 55 TO 85 | 1.010702 | 0.010510 | 0.644403 | -0.278089 | 1.052290 | -0.020025 | 106402. |
| CENTRES | 1.000193 | | 0.922492 | | 1.072315 | | |

PANEL III: FITTED VALUES FOR THREE FITS, ACTUAL VALUES AND RESIDUALS. (FIT 1 THROUGH 30, 40, 50 AND 60; FIT 2 THROUGH 55, 65, 75 AND 85; FIT 3 THROUGH 40, 50, 60 AND 70).

| AGE | FIT1 | FIT2 | FIT3 | ACTUAL | DEL1 | DEL2 | DEL3 |
|---|---|---|---|---|---|---|---|
| 20 | 83482. | 68089. | 83888. | 82983. | 498. | -14894. | 904. |
| 25 | 82155. | 70289. | 82413. | 81923. | 232. | -11633. | 490. |
| 30 | 80715. | 72111. | 80853. | 80715. | -0. | -8604. | 138. |
| 35 | 79099. | 73388. | 79150. | 79093. | 6. | -5706. | 57. |
| 40 | 77220. | 73918. | 77220. | 77220. | -0. | -3301. | -0. |
| 45 | 74947. | 73465. | 74932. | 74952. | -5. | -1487. | -20. |
| 50 | 72100. | 71766. | 72100. | 72100. | -0. | -333. | -0. |
| 55 | 68435. | 68565. | 68455. | 68565. | -131. | -0. | -110. |
| 60 | 63644. | 63654. | 63644. | 63644. | 0. | 10. | -0. |
| 65 | 57382. | 56947. | 57246. | 56947. | 435. | -0. | 299. |
| 70 | 49357. | 48570. | 48877. | 48877. | 481. | -307. | -0. |
| 75 | 39526. | 38949. | 38444. | 38949. | 577. | -0. | -506. |
| 80 | 28396. | 28847. | 26572. | 29128. | -732. | -282. | -2557. |
| 85 | 17300. | 19281. | 14983. | 19281. | -1981. | -0. | -4298. |
| 90 | 8205. | 11289. | 6131. | | | | |
| 95 | 2660. | 5572. | 1514. | | | | |
| 100 | 484. | 2206. | 169. | | | | |

on widowhood and even sibling survivorship. None of these sources can provide a complete life-table without resort to model life-tables, and indeed all use existing model life-tables in their derivation. However such estimates are a useful source of information on levels of mortality and can provide invaluable cross-checks on the accuracy of registration based life-tables. In addition, but less securely, such estimates can be used to adjust registration-based life-tables or segments thereof. In particular, in view of the lack of success of the approach through Hogan's model for assessing childhood mortality, there is a strong need for quasi-independent checks on the level of childhood mortality, especially as there is little reason for presuming childhood and adult mortality to be equally underregistered (see e.g. Table III.17 again). One quite substantial difficulty which arises in using childhood survival estimates to adjust registration based life-tables is that the most secure estimate obtained is $l_2$, although $l_3$ and, to a lesser extent, $l_5$ are also reasonably secure in most cases. Often, with registration data, deaths or death rates are only calculable for ages 0-1 and 1-4, giving only values of $l_1$ and $l_5$ for adjustment. As mentioned in the section on the Hogan model, fairly extensive trials were carried out to find a reasonable, simple functional relationship which would permit easy adjustment of $l_1$ and $l_5$ from an $l_2$ value (for example). These trials generally proved unsuccessful. An alternative, but less satisfactory solution is to use the estimate of $l_5$ from childhood survival to adjust $l_5$. The actual approach we have adopted is to use the ratio of the observed life-table ${}_5q_0$ to the childhood survival based ${}_5q_0$ as an adjustment factor for the death rates. This does not, in fact, lead to an adjusted value for ${}_5q_0$ which is the same as the childhood survival based one, but a slightly lower value. In view of the approximations involved, of the tendency for the childhood survival technique to give higher values of mortality than other approaches (see e.g. Brass (1971), who makes a different case) and of the childhood survival based estimates referring to an earlier period, a slight conservatism in the adjustment does not seem out of place.

When data from registration and the census are available by single years of age up to age 5 it is possible to do slightly better. But this involves placing some trust in the relative values of ${}_2q_0$, ${}_3q_0$ and ${}_5q_0$ (and perhaps, but very questionably ${}_1q_0$); this is not normally done and is very suspect. Nevertheless such an approach can sometimes prove useful, especially in providing a pointer to patterns of underregistration within the age group zero to five. We give an illustrative example for El Salvador, 1971 at Table III.21.

Table III.21. Comparison of registration and childhood survival based estimates of mortality, El-Salvador, 1971, Both Sexes.

| Age(x) | Registered ${}_xq_0$ | Childhood survival ${}_xq_0$ | Ratio | Registered ${}_xq_1$ | Childhood survival ${}_xq_1$ | Ratio |
|---|---|---|---|---|---|---|
| 1 | 0.06588 | 0.12864 | .512 | - | - | - |
| 2 | 0.08800 | 0.15140 | .581 | .02368 | .02612 | .907 |
| 3 | 0.09730 | 0.15160 | .642 | .03364 | .02635 | 1.277 |
| 4 | 0.10291 | - | | .03964 | - | - |
| 5 | 0.10688 | 0.17586 | .608 | .04389 | .05419 | .810 |

Source: Childhood survival rates, Soliz Sánchez (1976).

This example will take the estimates from childhood survival data as far as can be hoped and perhaps too far. Note for instance that ${}_2q_0$ and ${}_3q_0$ are almost identical, and that any attempt to infer, ${}_1q_0$ is doomed to failure. Deapite these reservations, it does seem reasonably clear that infant deaths are less well registered than those for ages 1 to 4. A fair approximation may

well be that around 51% of infant deaths are registered and around 81% of deaths in the range 1 to 4. However these relative estimates rely too heavily on the childhood survival based estimate of $_1q_0$ and a comparison of the $_3q_0$ values suggests that about 72% of deaths are registered in this range. (Note that we are quite loosely referring to the ratios of the q-values from the two sources as estimates of registration completeness: they are only approximately so and always tend to slightly underestimate such completeness owing to the non-linear relationship between q-values and m-values). This compares with around 58% for the ratios of the $_2q_0$'s. As childhood survival based estimate for $_2q_0$ is generally more reliable than that of $_1q_0$, the second pair of completeness estimates is used as the basis for correction. A major reservation in so doing is that infant mortality and mortality at age one are not likely to be equally underregistered. In view of this, the final approximation for El Salvador 1971 takes the level of 72% completeness as referring to ages 1 to 4, the level of 58% completeness as referring to ages 0-1; the implied completeness factor for infancy is 55%. The derivation and a comparison of the effects of the various estimates on the numbers of deaths are presented in Table III.22. Again, note that q-based

Table III.22. Effects of various completeness estimates on numbers of deaths in childhood, El Salvador 1971, both sexes.

| Age | Registered deaths | Completeness | Estimates | Completeness | Estimates | Completeness | Estimates |
|---|---|---|---|---|---|---|---|
| 0 | 8893 | 0.51 | 17,437 | 0.58 | 15,333 | 0.55 | 16,169 |
| 1 | 2544 | 0.81 | 3,141 | 0.58 | 4,386 | 0.72 | 3,533 |
| 2 | 1178 | 0.81 | 1,454 | 0.72 | 1,636 | 0.72 | 1,636 |
| 3 | 775 | 0.81 | 957 | 0.72 | 1,076 | 0.72 | 1,076 |
| 4 | 549 | 0.81 | 678 | 0.72 | 762 | 0.72 | 762 |
| Total | 13,939 | (.589) | 23,667 | (.601) | 23,193 | (.601) | 23,176 |

ratios are applied to deaths and thus to m-values. Note also, though, that the revised estimates of numbers of deaths would at best still be deficient by the same proportions as the census enumeration: as the project is to produce life-tables this presents little problem.

All three sets of completeness factors presented in Table III.22 lead to approximately the same total number of deaths, which is hardly surprising. However there is still some uncertainty about the pattern of mortality within childhood, even if the overall level is acceptable: infant mortality ($_1M_0$) varies between 119.6 $^0/_{00}$ and 136 $^0/_{00}$, $_1M_1$ ranges between 29.6$^0/_{00}$ and 41.4$^0/_{00}$. The last set of completeness estimates, which are also the median set, seem preferable.

8. Discussion

Several other problems must be mentioned. Many of the available data sets do not have death rates by single years of age up to age 5 and it thus becomes virtually impossible to make any inferences about difference in registration completeness, within childhood, with only $_1l_1$ and $_1l_5$ available. Overall completeness can be assessed from a value of $_1l_5$, but it should be noted that this value from childhood survival data is not the most secure: an assessment has to be made of the consistency and trustworthiness of the $_nq_0$'s from childhood survival before use of $_1l_5$ ($_5q_0$) for such adjustment. In cases where the childhood survival data appear highly consistent, it can prove possible to use the estimate of $_1q_0$ to derive a very approximate estimate of underregistration of infant deaths. Such an estimate should always be treated very cautiously. At least some assessment of overall completeness levels in childhood is better than none at all, even if infant deaths are expected to be differentially underregistered.

The next point is that, even with single year registration data, there is considerable doubt about the implied pattern of mortality within early childhood. Such uncertainty would rule out except within broad limits the possibility of distinguishing 'families' of model tables on the basis of patterns within childhood, although the relationship between child and adult

mortality could still provide a means of identifying 'families'.

Finally, it is necessary to look at the implications of the results for 1971 on the data for El Salvador. Estimates of completeness of childhood mortality are only available for both sexes combined. As no strong evidence of sex differentials in completeness of registration for adults was found it will be assumed that this is also the case within childhood, although with lower completeness levels this must be subject to doubt. Recall our estimated level of completeness for adult registration in 1971, which was 96%. Thus there appears to be substantial differential under-registration. Note that these differentials are not out of line with those shown for Honduras in Table III.17, which are much more securely based. The remaining problem for El Salvador is how to estimate completeness of childhood registration for earlier periods. No childhood survival data exist for these earlier periods. The estimates of completeness shown at Table III.23 have been adopted. These estimates are based on a pro-rata adjustment of the

Table III.23. Estimates of completeness of death registration for El Salvador

| | Age range | | |
|---|---|---|---|
| Year | 0 | 1-4 | 5 and over |
| 1971 | .58 | .72 | .96 |
| 1961 | .53 | .66 | .88 |
| 1950 | .48 | .60 | .80 |

1971 childhood completeness estimates by the ratio of the adult completeness estimates. Although this is of doubtful validity it probably errs on the cautious side, as it is clearly harder to progress from 88% to 96% completeness than from 53% to 58%. Thus it may well be that childhood mortality was even

less well registered in the earlier periods than is implied by these estimates.

Although the estimates of differential childhood underregistration of deaths are based on quite limited information, the resulting estimates are almost certainly to be preferred over simply using the adult registration estimates supplied to all ages or the raw registration data.

9. Calculation of life-tables

Whether or not the registration data are adjusted, it is necessary to derive life-tables therefrom. Life-tables are also quite heavily used as inputs to some of the assessment and correction procedures. As there are always minor variations in the methods used for life-table construction it is important to spell out the assumptions made. There are two crucial areas where assumptions are necessary. The first is the relationship between ${}_nM_x$ (the observed death rate) and ${}_nm_x$ (the life-table death rate): this relationship is determined by the interaction of the age-structure and the force of mortality, $\mu(x)$, as

$$ {}_nM_x = \frac{\int_0^n P(x+t)\,\mu(x+t)\,dt}{\int_0^n P(x+t)\,dt} $$

and

$$ {}_nm_x = \frac{\int_0^n l(x+t)\,\mu(x+t)\,dt}{\int_0^n l(x+t)\,dt.} $$

If the population age-structure is different from the life-table age-structure and $\mu(x)$ is changing within the group, the two measures will typically not be identical. Owing to the difficulty of overcoming this non-equality, it is common to assume that equality of these two measures exists (or to ignore or fail to recognize existence of a difference). For lack of a suitable alternative for ages up to ten, equality is assumed, although it probably does not truly exist. For the higher ages this assumption is avoided by using an approximation due to Keyfitz and Frauenthal (1975), which also implies an assumption about the second crucial element in life-table construction, namely the relationship between ${}_nm_x$ and ${}_np_x$ or ${}_nq_x$. This can always be written as

$$ {}_nq_x = \frac{n \cdot {}_nm_x}{1-(n-{}_na_x)\ {}_nm_x} $$

where ${}_na_x$ is the average number of years lived within the age interval x to x+n by those dying within the interval. Thus the second crucial assumption is the value of ${}_na_x$, whether derived implicitly or defined explicitly. A particular value of ${}_na_x$ also specifies the value of ${}_nL_x$ in terms of $l_x$ and $l_{x+n}$, namely ${}_nL_x = {}_na_x \cdot l_x + (n-{}_na_x)\, l_{x+n}$.

The set of assumptions used here are now outlined.

(i) Ages under 10

${}_nM_x = {}_nm_x$ for all ages under 10 and ${}_5a_5 = 2.5$.

For ages under five, the relationships used by Coale and Demeny (1966) for their West Model Life Tables were adapted. The adaptations involve converting their formulae for ${}_1a_0$ and ${}_4a_1$ from expressions in ${}_1q_0$ to expressions in ${}_1M_0$. Secondly, we changed the assumed values of ${}_1a_1$, ${}_1a_2$, ${}_1a_3$ and ${}_1a_4$ to 0.44, 0.47, 0.49 and 0.50 respectively, instead of a constant value of 0.50; these values were used directly when data were available by single year of age from 1 to 4. The expressions used were:

| | | Males | Females | Both sexes |
|---|---|---|---|---|
| for ${}_1M_0 < 0.107$ | ${}_1a_0$ | $0.0425 + 2.687\,{}_1M_0$ | $0.050 + 2.804\,{}_1M_0$ | $0.0463 + 2.745\,{}_1M_0$ |
| | ${}_4a_1$ | $1.621 - 2.879\,{}_1M_0$ | $1.488 - 1.554\,{}_1M_0$ | $1.554 - 2.206\,{}_1M_0$ |
| for ${}_1M_0 \geq 0.107$ | ${}_1a_0$ | 0.33 | 0.35 | 0.34 |
| | ${}_4a_1$ | 1.313 | 1.322 | 1.318 |

(ii) Ages 10 and over

The approximation derived by Keyfitz and Frauenthal (see also Keyfitz (1977)) presumes linearity both of the population and of the force of mortality over successive segments of the age-range. This approximation is convenient for present purposes. The approximations used were:

$$ {}_5P_x = \exp\left[-5\,{}_5M_x + \frac{5}{48\,{}_5P_x}\,({}_5P_{x-5} - {}_5P_{x+5})\,({}_5M_{x+5} - {}_5M_{x-5})\right] $$

$$ = \exp\left[-5({}_5M_x + C_x)\right], $$

$$ \text{where } C_x = \frac{({}_5P_{x-5} - {}_5P_{x+5})\,({}_5M_{x+5} - {}_5M_{x-5})}{48\,{}_5P_x} $$

$$ \text{Then } {}_5L_x = \frac{5(l_x - l_{x+5})}{\ln(l_x) - \ln(l_{x+5})}\left[1 + \frac{5}{24}\left[{}_5M_{x+5} - {}_5M_{x-5}\right]\right] $$

and

$$ {}_5m_x = \frac{{}_5M_x + C_x}{\left[1 + \frac{5}{24}\,({}_5M_{x+5} - {}_5M_{x-5})\right]} $$

also denoting $B_x = 5/24\,({}_5M_{x+5} - {}_5M_{x-5})$.

A final slight difficulty arises from the fact that ${}_5M_{x+5}$ and ${}_5P_{x+5}$ are not available for the last five-year group or for the open-ended group. Denoting the penultimate five-year age-group as starting at age L, so that the last available values of $C_x$ and $B_x$, which entirely specify both the relationship between ${}_nM_x$ and ${}_nm_x$ and the value of ${}_na_x$, are $C_L$ and $B_L$. After some experimentation it appeared quite reasonable to extrapolate these values to the last two age groups using ratios of previous terms. In practice the following relationships were used to complete the life-tables:

$$B_{L+5} = B_L \ (B_L / B_{L-20})^{0.25}$$

$$C_{L+5} = C_L \ (C_L / C_{L-20})^{0.25}$$

$$B_{L+10} = B_{L+5} \ (B_L / B_{L-20})^{\left[(85-L)/10\right]}$$

$$C_{L+10} = C_{L+5} \ (C_L / C_{L-20})^{\left[(85-L)/10\right]}$$

The exponent of (85-L/10 is perhaps slightly conservative, but as extrapolation is involved it is wise to reduce the chances of values of $C_x$ and $B_x$ becoming too large.

Table III.24 shows examples of a life-table calculated by this method for El Salvador Females, 1971. A 13.32% increase in the numbers of deaths only leads to a 4.9% decrease in $\mathring{e}_o$. This kind of relationship has recently been extensively studied by Keyfitz (1977 and several papers).

It is now necessary to comment on the reasonableness of the assumptions. As already indicated, the assumption of equality of ${}_nM_x$ and ${}_nm_x$ up to age ten is not ideal, especially in the first few years of life, although this will not introduce any great bias, but the observed death rates will be higher than the true life-table rates in a growing population. It is also

TABLE III.24. LIFE TABLES FOR EL SALVADOR, FEMALES, 1961 BASED ON UNCORRECTED REGISTRATION DATA.

| AGE | $_nP_x$ | $_nq_x$ | $l_x$ | $_nd_x$ | $_na_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.92016 | 0.07984 | 100000. | 7984. | 0.28737 | 94311. | 5978919. | 59.79 | 0.08465 | 0.08465 | | |
| 1 | 0.96675 | 0.03325 | 92016. | 3060. | 0.44000 | 90303. | 5884608. | 63.95 | 0.03388 | 0.03388 | | |
| 2 | 0.98228 | 0.01772 | 88957. | 1577. | 0.47000 | 88121. | 5794305. | 65.14 | 0.01789 | 0.01789 | | |
| 3 | 0.98901 | 0.01099 | 87380. | 960. | 0.49000 | 86890. | 5706184. | 65.30 | 0.01105 | 0.01105 | | |
| 4 | 0.99326 | 0.00674 | 86420. | 583. | 0.50000 | 86129. | 5619294. | 65.02 | 0.00676 | 0.00676 | | |
| 5 | 0.98409 | 0.01591 | 85837. | 1365. | 2.49332 | 425764. | 5533165. | 64.46 | 0.00321 | 0.00321 | | |
| 10 | 0.99293 | 0.00707 | 84472. | 598. | 2.33887 | 420769. | 5107401. | 60.46 | 0.00142 | 0.00143 | -0.00001 | -0.00022 |
| 15 | 0.98938 | 0.01062 | 83874. | 891. | 2.60658 | 417239. | 4686632. | 55.88 | 0.00214 | 0.00213 | 0.00001 | 0.00024 |
| 20 | 0.98722 | 0.01278 | 82983. | 1061. | 2.56187 | 412331. | 4269392. | 51.45 | 0.00257 | 0.00257 | 0.00001 | 0.00017 |
| 25 | 0.98525 | 0.01475 | 81923. | 1208. | 2.59763 | 406711. | 3857062. | 47.08 | 0.00297 | 0.00296 | 0.00001 | 0.00031 |
| 30 | 0.97992 | 0.02008 | 80715. | 1621. | 2.58522 | 399658. | 3450351. | 42.75 | 0.00406 | 0.00405 | 0.00001 | 0.00038 |
| 35 | 0.97631 | 0.02369 | 79093. | 1874. | 2.57187 | 390917. | 3050693. | 38.57 | 0.00479 | 0.00478 | 0.00001 | 0.00039 |
| 40 | 0.97063 | 0.02937 | 77220. | 2268. | 2.59040 | 380634. | 2659776. | 34.44 | 0.00596 | 0.00593 | 0.00003 | 0.00061 |
| 45 | 0.96195 | 0.03805 | 74952. | 2852. | 2.59238 | 367892. | 2279142. | 30.41 | 0.00775 | 0.00773 | 0.00003 | 0.00084 |
| 50 | 0.95098 | 0.04902 | 72100. | 3534. | 2.62580 | 352107. | 1911249. | 26.51 | 0.01004 | 0.00998 | 0.00008 | 0.00148 |
| 55 | 0.92822 | 0.07178 | 68565. | 4921. | 2.63876 | 331207. | 1559142. | 22.74 | 0.01486 | 0.01481 | 0.00009 | 0.00253 |
| 60 | 0.89477 | 0.10523 | 63644. | 6697. | 2.59592 | 302120. | 1227935. | 19.29 | 0.02217 | 0.02212 | 0.00012 | 0.00316 |
| 65 | 0.85929 | 0.14171 | 56947. | 8070. | 2.59232 | 265305. | 925815. | 16.26 | 0.03042 | 0.02999 | 0.00058 | 0.00477 |
| 70 | 0.79689 | 0.20311 | 48877. | 9927. | 2.53132 | 219876. | 660510. | 13.51 | 0.04515 | 0.04500 | 0.00041 | 0.00571 |
| 75 | 0.74785 | 0.25215 | 38949. | 9821. | 2.50947 | 170287. | 440634. | 11.31 | 0.05767 | 0.05742 | 0.00070 | 0.00758 |
| 80 | 0.66193 | 0.33807 | 29128. | 9847. | 2.44935 | 120524. | 270347. | 9.28 | 0.08170 | 0.08136 | 0.00116 | 0.00997 |
| 85 | 0.0 | 1.00000 | 19281. | 19281. | 7.77054 | 149823. | 149823. | 7.77 | 0.12869 | 0.12845 | 0.00193 | 0.01311 |

assumed that $_1q_0$ can be derived from $_1m_0$ in a simple way. In general it is preferable (because of the non-equality of $_1M_0$ and $_1m_0$) to calculate $_1q_0$ directly from detailed information on births and deaths by sub-intervals of age within the first year of life for a two year period. Often this has proved impossible. There are also problems introduced by differential registration of births and infant deaths. Where possible we have compared the two to check the validity of the calculations, which are themselves susceptible to differences in enumeration and registration completeness. Where possible these have been, as noted earlier, checked against other sources such as childhood survival based estimates. The choice of values used for the West model life-tables for the approximations to $_1a_0$ and $_4a_1$ should not cause any serious biasses. The values and relationships used for the other regional sets were not very different and the range of choice would make very little impact on the resulting life-table. The other set of questions arise about the approximation used for the higher ages. Obviously neither $\mu(x)$ nor $P(x)$ are linear over five-year ranges and the method of approximation to the linear effects actually uses a fifteen year range. In particular no use is made of the central five-year portion of the fifteen year range to estimate these slopes. This is suspect, but, in general and especially for developing countries which usually have monotonically declining age-structures, is almost certainly slightly more accurate than ignoring the differences between $_nM_x$ and $_nm_x$. In particular, Keyfitz and Frauenthal (1975) demonstrate that this approach subsumes Reed and Merrell's and Greville's approximations as special cases.

10. Adjusted life-tables for El Salvador

As a result of the evaluation and adjustment procedures applied to the data for El Salvador throughout this section it is possible to produce adjusted life-tables which can be used as input for the model life-table project. Table III.23 gave completeness estimates for the various age-segments and time-periods. These have been used to produce the adjusted tables. Table III.25 shows summary indicators of the unadjusted and adjusted life-tables for El Salvador. Tables III.26 to III.28 present the detailed life-tables for 1950, 1961 and 1971 respectively, for both males and females. Adjusted life-tables for both sexes are also available but are not presented here. These adjusted tables are in reasonable agreement with other

Table III.25. Comparison of unadjusted and adjusted expectations of life at birth and at age 5 for El Salvador, 1950, 1961 and 1971

| Year | Sex | Expectation of life at birth | | Expectation of life at age 5 | |
|---|---|---|---|---|---|
| | | Unadjusted | Adjusted | Unadjusted | Adjusted |
| 1950 | M | 49.06 | 37.96 | 56.11 | 51.93 |
| 1950 | F | 51.59 | 40.73 | 58.11 | 54.00 |
| 1961 | M | 55.77 | 47.80 | 61.77 | 58.57 |
| 1961 | F | 59.79 | 52.12 | 65.14 | 61.95 |
| 1971 | M | 59.39 | 54.81 | 63.58 | 62.43 |
| 1971 | F | 64.54 | 60.25 | 68.16 | 67.03 |

estimates. Bixby (1976) calculates life-tables for El Salvador on the basis of inter-censal survivorship ratios which give $\mathring{e}_5$ as 60.67 for males and 62.47 for females. The results presented here also agree quite well with the estimates given by the United Nations in World Population Prospects as Assessed in 1973 and shown below.

Table III.25a. Estimates of expectation of life at birth for El Salvador from World Population Prospects as Assessed in 1973 (U.N. 1977)

| | Year | | | | |
|---|---|---|---|---|---|
| | 50/55 | 55/60 | 60/65 | 65/70 | 70/75 |
| Males | 42.6 | 45.7 | 50.0 | 53.2 | 56.0 |
| Females | 44.9 | 48.6 | 53.2 | 56.7 | 59.7 |

In general the adjusted estimates are compatible with other sources, although suggesting lower expectations of life at birth in earlier years. The estimates and adjustments presented here for El Salvador are to be preferred to other available estimates as they are based on a fairly extensive analysis using modern techniques.

TABLE III.26. ADJUSTED LIFE-TABLES FOR EL SALVADOR, 1950.

(SCALING OF DEATHS; INFANCY BY 2.0830, AGES 1 TO 4 BY 1.6670 AND 5 AND OVER BY 1.2500)

MALES

| AGE | $_nP_x$ | $_nq_x$ | $l_x$ | $_nd_x$ | $_na_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.79377 | 0.20623 | 100000. | 20623. | 0.29427 | 85446. | 3796270. | 37.96 | 0.24136 | 0.24136 | | |
| 1 | 0.83160 | 0.16820 | 79377. | 13351. | 1.35124 | 282144. | 3710824. | 46.75 | 0.04732 | 0.04732 | | |
| 5 | 0.96165 | 0.03835 | 66026. | 2532. | 2.48371 | 323759. | 3428680. | 51.93 | 0.00782 | 0.00782 | | |
| 10 | 0.98429 | 0.01571 | 63494. | 997. | 2.33284 | 314810. | 3104922. | 48.90 | 0.00317 | 0.00318 | -0.00001 | -0.00051 |
| 15 | 0.97328 | 0.02672 | 62497. | 1670. | 2.69583 | 308635. | 2790112. | 44.64 | 0.00541 | 0.00538 | 0.00004 | 0.00112 |
| 20 | 0.95793 | 0.04207 | 60826. | 2559. | 2.57787 | 297935. | 2481477. | 40.80 | 0.00859 | 0.00857 | 0.00003 | 0.00082 |
| 25 | 0.95435 | 0.04565 | 58268. | 2660. | 2.51538 | 284730. | 2183542. | 37.47 | 0.00934 | 0.00933 | 0.00001 | 0.00033 |
| 30 | 0.95058 | 0.04942 | 55608. | 2748. | 2.51412 | 271207. | 1898813. | 34.15 | 0.01013 | 0.01013 | 0.00001 | 0.00036 |
| 35 | 0.94620 | 0.05380 | 52860. | 2844. | 2.54573 | 257318. | 1627605. | 30.79 | 0.01105 | 0.01105 | 0.00001 | 0.00076 |
| 40 | 0.93320 | 0.06680 | 50016. | 3341. | 2.54709 | 241883. | 1370287. | 27.40 | 0.01381 | 0.01378 | 0.00005 | 0.00105 |
| 45 | 0.92250 | 0.07750 | 46675. | 3617. | 2.54796 | 224503. | 1128404. | 24.18 | 0.01611 | 0.01608 | 0.00005 | 0.00132 |
| 50 | 0.90398 | 0.09602 | 43057. | 4134. | 2.54164 | 205123. | 903901. | 20.99 | 0.02016 | 0.02010 | 0.00009 | 0.00169 |
| 55 | 0.88511 | 0.11489 | 38923. | 4472. | 2.57474 | 183770. | 698778. | 17.95 | 0.02433 | 0.02419 | 0.00021 | 0.00306 |
| 60 | 0.83959 | 0.16042 | 34451. | 5527. | 2.54683 | 158699. | 515009. | 14.95 | 0.03482 | 0.03481 | 0.00016 | 0.00418 |
| 65 | 0.79764 | 0.20236 | 28925. | 5853. | 2.58653 | 130497. | 356310. | 12.32 | 0.04485 | 0.04428 | 0.00094 | 0.00817 |
| 70 | 0.68763 | 0.31237 | 23071. | 7207. | 2.48053 | 97200. | 225813. | 9.79 | 0.07414 | 0.07402 | 0.00088 | 0.01020 |
| 75 | 0.62433 | 0.37567 | 15865. | 5960. | 2.40820 | 63876. | 128614. | 8.11 | 0.09330 | 0.09325 | 0.00097 | 0.00978 |
| 80 | 0.54238 | 0.45762 | 9905. | 4533. | 2.35343 | 37528. | 64737. | 6.54 | 0.12078 | 0.12095 | 0.00141 | 0.01306 |
| 85 | 0.0 | 1.00000 | 5372. | 5372. | 5.06495 | 27210. | 27210. | 5.06 | 0.19744 | 0.19882 | 0.00206 | 0.01746 |

FEMALES

| AGE | $_nP_x$ | $_nq_x$ | $l_x$ | $_nd_x$ | $_na_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.81583 | 0.18417 | 100000. | 18417. | 0.31273 | 87343. | 4072583. | 40.73 | 0.21086 | 0.21086 | | |
| 1 | 0.83848 | 0.16152 | 81583. | 13177. | 1.34239 | 291312. | 3985240. | 48.85 | 0.04523 | 0.04523 | | |
| 5 | 0.96005 | 0.03995 | 68406. | 2733. | 2.48301 | 335150. | 3693928. | 54.00 | 0.00815 | 0.00815 | | |
| 10 | 0.98522 | 0.01478 | 65673. | 971. | 2.24415 | 325688. | 3358778. | 51.14 | 0.00298 | 0.00299 | -0.00002 | -0.00074 |
| 15 | 0.97729 | 0.02271 | 64702. | 1470. | 2.63511 | 320034. | 3033090. | 46.88 | 0.00459 | 0.00459 | 0.00001 | 0.00066 |
| 20 | 0.96946 | 0.03054 | 63232. | 1931. | 2.58519 | 311499. | 2713055. | 42.91 | 0.00620 | 0.00618 | 0.00002 | 0.00061 |
| 25 | 0.96298 | 0.03702 | 61301. | 2269. | 2.58569 | 301028. | 2401557. | 39.18 | 0.00754 | 0.00751 | 0.00004 | 0.00077 |
| 30 | 0.95185 | 0.04815 | 59032. | 2842. | 2.52096 | 288114. | 2100529. | 35.58 | 0.00986 | 0.00986 | 0.00001 | 0.00041 |
| 35 | 0.95368 | 0.04632 | 56190. | 2603. | 2.53827 | 274543. | 1812414. | 32.26 | 0.00948 | 0.00947 | 0.00001 | 0.00055 |
| 40 | 0.93924 | 0.06076 | 53587. | 3256. | 2.53487 | 259909. | 1537872. | 28.70 | 0.01253 | 0.01250 | 0.00004 | 0.00076 |
| 45 | 0.93626 | 0.06374 | 50331. | 3208. | 2.53203 | 243737. | 1277963. | 25.39 | 0.01316 | 0.01314 | 0.00003 | 0.00078 |
| 50 | 0.92157 | 0.07843 | 47123. | 3696. | 2.55805 | 226589. | 1034225. | 21.95 | 0.01631 | 0.01626 | 0.00008 | 0.00150 |
| 55 | 0.90190 | 0.09810 | 43427. | 4260. | 2.65145 | 207130. | 807636. | 18.60 | 0.02057 | 0.02036 | 0.00029 | 0.00402 |
| 60 | 0.83633 | 0.16367 | 39167. | 6411. | 2.56016 | 180193. | 600506. | 15.33 | 0.03558 | 0.03554 | 0.00021 | 0.00481 |
| 65 | 0.80178 | 0.19822 | 32756. | 6493. | 2.55350 | 147896. | 420313. | 12.83 | 0.04390 | 0.04346 | 0.00072 | 0.00643 |
| 70 | 0.71514 | 0.28486 | 26263. | 7481. | 2.48915 | 112533. | 272417. | 10.37 | 0.06648 | 0.06639 | 0.00067 | 0.00862 |
| 75 | 0.65156 | 0.34844 | 18782. | 6545. | 2.42769 | 77076. | 159884. | 8.51 | 0.08491 | 0.08485 | 0.00083 | 0.00905 |
| 80 | 0.57430 | 0.42570 | 12238. | 5209. | 2.37008 | 47487. | 82808. | 6.77 | 0.10970 | 0.10983 | 0.00109 | 0.01109 |
| 85 | 0.0 | 1.00000 | 7028. | 7028. | 5.02567 | 35321. | 35321. | 5.03 | 0.19898 | 0.20026 | 0.00142 | 0.01359 |

TABLE III.27. ADJUSTED LIFE-TABLES FOR EL SALVADOR, 1961.
(SCALING OF DEATHS; INFANCY BY 1.8870, AGES 1 TO 4 BY 1.5150, AND 5 AND OVER BY 1.1560)

MALES

| AGE | $_nP_x$ | $_nQ_x$ | $l_x$ | $_nd_x$ | $_nm_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.83002 | 0.16998 | 100000. | 16998. | 0.29427 | 88004. | 4779516. | 47.80 | 0.19315 | 0.19315 | | |
| 1 | 0.94841 | 0.05159 | 83002. | 4282. | 0.44000 | 80604. | 4691512. | 56.52 | 0.05312 | 0.05312 | | |
| 2 | 0.97301 | 0.02699 | 78720. | 2125. | 0.47000 | 77594. | 4610908. | 58.57 | 0.02738 | 0.02738 | | |
| 3 | 0.98271 | 0.01729 | 76595. | 1325. | 0.49000 | 75920. | 4533314. | 59.19 | 0.01745 | 0.01745 | | |
| 4 | 0.98953 | 0.01047 | 75271. | 788. | 0.50000 | 74877. | 4457395. | 59.22 | 0.01052 | 0.01052 | | |
| 5 | 0.98025 | 0.01975 | 74483. | 1471. | 2.49169 | 368724. | 4382518. | 58.84 | 0.00399 | 0.00399 | | |
| 10 | 0.99144 | 0.00856 | 73012. | 625. | 2.37311 | 363417. | 4013794. | 54.97 | 0.00172 | 0.00173 | -0.00001 | -0.00021 |
| 15 | 0.98507 | 0.01493 | 72387. | 1080. | 2.72883 | 359481. | 3650377. | 50.43 | 0.00301 | 0.00297 | 0.00004 | 0.00071 |
| 20 | 0.97459 | 0.02541 | 71306. | 1812. | 2.61915 | 352219. | 3290896. | 46.15 | 0.00514 | 0.00512 | 0.00002 | 0.00067 |
| 25 | 0.96951 | 0.03049 | 69495. | 2119. | 2.53811 | 342259. | 2938677. | 42.29 | 0.00619 | 0.00618 | 0.00001 | 0.00032 |
| 30 | 0.96733 | 0.03267 | 67376. | 2201. | 2.50932 | 331400. | 2596418. | 38.54 | 0.00664 | 0.00664 | 0.00000 | 0.00015 |
| 35 | 0.96594 | 0.03406 | 65175. | 2220. | 2.54098 | 320419. | 2265019. | 34.75 | 0.00693 | 0.00692 | 0.00001 | 0.00038 |
| 40 | 0.95832 | 0.04168 | 62956. | 2624. | 2.57731 | 308421. | 1944600. | 30.89 | 0.00851 | 0.00848 | 0.00003 | 0.00081 |
| 45 | 0.94718 | 0.05282 | 60332. | 3186. | 2.58699 | 293969. | 1636178. | 27.12 | 0.01084 | 0.01080 | 0.00005 | 0.00119 |
| 50 | 0.93104 | 0.06896 | 57145. | 3941. | 2.60759 | 276297. | 1342210. | 23.49 | 0.01426 | 0.01419 | 0.00010 | 0.00196 |
| 55 | 0.90344 | 0.09656 | 53204. | 5137. | 2.57580 | 253567. | 1065913. | 20.03 | 0.02026 | 0.02023 | 0.00008 | 0.00240 |
| 60 | 0.87876 | 0.12124 | 48067. | 5827. | 2.59683 | 226329. | 812346. | 16.90 | 0.02575 | 0.02570 | 0.00014 | 0.00389 |
| 65 | 0.81976 | 0.18024 | 42239. | 7613. | 2.57692 | 192749. | 586017. | 13.87 | 0.03950 | 0.03892 | 0.00083 | 0.00635 |
| 70 | 0.75283 | 0.24717 | 34626. | 8559. | 2.52146 | 151917. | 393268. | 11.36 | 0.05634 | 0.05617 | 0.00061 | 0.00793 |
| 75 | 0.67745 | 0.32255 | 26067. | 8408. | 2.45217 | 108915. | 241350. | 9.26 | 0.07720 | 0.07697 | 0.00091 | 0.00888 |
| 80 | 0.60513 | 0.39487 | 17659. | 6973. | 2.41418 | 70266. | 132435. | 7.50 | 0.09924 | 0.09879 | 0.00167 | 0.01232 |
| 85 | 0.0 | 1.00000 | 10686. | 10686. | 5.81775 | 62170. | 62170. | 5.82 | 0.17189 | 0.17177 | 0.00306 | 0.01709 |

FEMALES

| AGE | $_nP_x$ | $_nQ_x$ | $l_x$ | $_nd_x$ | $_nm_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.85606 | 0.14394 | 100000. | 14394. | 0.31273 | 90108. | 5212408. | 52.12 | 0.15974 | 0.15974 | | |
| 1 | 0.95010 | 0.04990 | 85606. | 4271. | 0.44000 | 83214. | 5122300. | 59.84 | 0.05133 | 0.05133 | | |
| 2 | 0.97328 | 0.02672 | 81335. | 2174. | 0.47000 | 80183. | 5039086. | 61.95 | 0.02711 | 0.02711 | | |
| 3 | 0.98340 | 0.01660 | 79161. | 1314. | 0.49000 | 78491. | 4958904. | 62.64 | 0.01674 | 0.01674 | | |
| 4 | 0.98980 | 0.01020 | 77847. | 794. | 0.50000 | 77451. | 4880413. | 62.69 | 0.01025 | 0.01025 | | |
| 5 | 0.98195 | 0.01805 | 77054. | 1391. | 2.49241 | 381780. | 4802962. | 62.33 | 0.00364 | 0.00364 | | |
| 10 | 0.99197 | 0.00803 | 75663. | 608. | 2.33847 | 376696. | 4421182. | 58.43 | 0.00161 | 0.00162 | -0.00001 | -0 00026 |
| 15 | 0.98794 | 0.01206 | 75055. | 905. | 2.60597 | 373108. | 4044485. | 53.89 | 0.00243 | 0.00242 | 0.00001 | 00027 |
| 20 | 0.98549 | 0.01451 | 74150. | 1076. | 2.56114 | 368126. | 3671377. | 49.51 | 0.00292 | 0.00292 | 0.00001 | 0.00020 |
| 25 | 0.98326 | 0.01674 | 73074. | 1223. | 2.59679 | 362432. | 3303251. | 45.20 | 0.00337 | 0.00336 | 0.00001 | 0.00035 |
| 30 | 0.97722 | 0.02278 | 71851. | 1637. | 2.58407 | 355301. | 2940820. | 40.93 | 0.00461 | 0.00460 | 0.00001 | 0.00043 |
| 35 | 0.97313 | 0.02687 | 70214. | 1887. | 2.57051 | 346487. | 2585519. | 36.82 | 0.00544 | 0.00543 | 0.00001 | 0.00045 |
| 40 | 0.96671 | 0.03329 | 68328. | 2275. | 2.58871 | 336152. | 2239032. | 32.77 | 0.00677 | 0.00674 | 0.00003 | 0.00070 |
| 45 | 0.95688 | 0.04312 | 66053. | 2848. | 2.59018 | 323400. | 1902880. | 28.81 | 0.00881 | 0.00878 | 0.00004 | 0.00096 |
| 50 | 0.94450 | 0.05550 | 63205. | 3508. | 2.62295 | 307685. | 1579480. | 24.99 | 0.01140 | 0.01133 | 0.00009 | 0.00168 |
| 55 | 0.91887 | 0.08113 | 59697. | 4843. | 2.63454 | 287028. | 1271794. | 21.30 | 0.01687 | 0.01682 | 0.00010 | 0.00287 |
| 60 | 0.88135 | 0.11865 | 54854. | 6509. | 2.58962 | 258580. | 984766. | 17.95 | 0.02517 | 0.02512 | 0.00014 | 0.00359 |
| 65 | 0.84063 | 0.15937 | 48345. | 7705. | 2.58367 | 223108. | 726186. | 15.02 | 0.03453 | 0.03407 | 0.00065 | 0.00542 |
| 70 | 0.77266 | 0.22734 | 40640. | 9239. | 2.51849 | 180275. | 503077. | 12.38 | 0.05125 | 0.05112 | 0.00047 | 0.00649 |
| 75 | 0.71887 | 0.28113 | 31401. | 8828. | 2.49309 | 134875. | 322803. | 10.28 | 0.06545 | 0.06522 | 0.00079 | 0.00861 |
| 80 | 0.62561 | 0.37419 | 22573. | 8447. | 2.42619 | 91127. | 187928. | 8.33 | 0.09269 | 0.09243 | 0.00132 | 0.01132 |
| 85 | 0.0 | 1.00000 | 14127. | 14127. | 6.85230 | 96801. | 96801. | 6.85 | 0.14594 | 0.14592 | 0.00219 | 0.01489 |

TABLE III.28. ADJUSTED LIFE-TABLES FOR EL SALVADOR, 1971.
(SCALING OF DEATHS; INFANCY BY 1.7240, AGES 1 TO 4 BY 1.3890, AND 5 AND OVER BY 1.0420)

MALES

| AGE | $_nP_x$ | $_nq_x$ | $l_x$ | $_nd_x$ | $_na_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.87923 | 0.12077 | 100000. | 12077. | 0.29427 | 91477. | 5481485. | 54.81 | 0.13202 | 0.13202 | | |
| 1 | 0.96625 | 0.03375 | 87923. | 2967. | 0.44000 | 86261. | 5390008. | 61.30 | 0.03440 | 0.03440 | | |
| 2 | 0.98539 | 0.01461 | 84956. | 1241. | 0.47000 | 84298. | 5303747. | 62.43 | 0.01473 | 0.01473 | | |
| 3 | 0.99151 | 0.00849 | 83715. | 710. | 0.49000 | 83352. | 5219449. | 62.35 | 0.00852 | 0.00852 | | |
| 4 | 0.99382 | 0.00618 | 83004. | 513. | 0.50000 | 82748. | 5136096. | 61.88 | 0.00620 | 0.00620 | | |
| 5 | 0.98683 | 0.01317 | 82491. | 1086. | 2.49448 | 409736. | 5053349. | 61.26 | 0.00265 | 0.00265 | | |
| 10 | 0.99364 | 0.00636 | 81405. | 518. | 2.38383 | 405671. | 4643613. | 57.04 | 0.00128 | 0.00128 | -0.00001 | -0.00014 |
| 15 | 0.99011 | 0.00989 | 80887. | 800. | 2.77605 | 402658. | 4237942. | 52.39 | 0.00199 | 0.00196 | 0.00003 | 0.00056 |
| 20 | 0.98029 | 0.01971 | 80088. | 1578. | 2.62689 | 396693. | 3835284. | 47.89 | 0.00398 | 0.00396 | 0.00002 | 0.00054 |
| 25 | 0.97752 | 0.02248 | 78509. | 1765. | 2.53739 | 388200. | 3438591. | 43.80 | 0.00455 | 0.00454 | 0.00001 | 0.00021 |
| 30 | 0.97534 | 0.02461 | 76744. | 1889. | 2.54231 | 379080. | 3050391. | 39.75 | 0.00498 | 0.00498 | 0.00000 | 0.00026 |
| 35 | 0.97135 | 0.02865 | 74856. | 2145. | 2.57779 | 369083. | 2671312. | 35.69 | 0.00581 | 0.00580 | 0.00001 | 0.00052 |
| 40 | 0.96310 | 0.03690 | 72711. | 2683. | 2.58545 | 357076. | 2302229. | 31.66 | 0.00751 | 0.00749 | 0.00003 | 0.00076 |
| 45 | 0.95362 | 0.04638 | 70028. | 3248. | 2.59703 | 342318. | 1945153. | 27.78 | 0.00949 | 0.00945 | 0.00005 | 0.00106 |
| 50 | 0.93856 | 0.06144 | 66780. | 4103. | 2.61891 | 324131. | 1602835. | 24.00 | 0.01266 | 0.01258 | 0.00010 | 0.00184 |
| 55 | 0.91210 | 0.08790 | 62677. | 5509. | 2.60015 | 300164. | 1278704. | 20.40 | 0.01835 | 0.01830 | 0.00010 | 0.00255 |
| 60 | 0.88259 | 0.11741 | 57168. | 6712. | 2.62017 | 269866. | 978540. | 17.12 | 0.02487 | 0.02482 | 0.00016 | 0.00430 |
| 65 | 0.82143 | 0.17857 | 50456. | 9010. | 2.55636 | 230263. | 708673. | 14.05 | 0.03913 | 0.03894 | 0.00040 | 0.00544 |
| 70 | 0.77329 | 0.22671 | 41446. | 9396. | 2.53806 | 184096. | 478411. | 11.54 | 0.05104 | 0.05093 | 0.00049 | 0.00746 |
| 75 | 0.68244 | 0.31756 | 32050. | 10178. | 2.49188 | 134721. | 294314. | 9.18 | 0.07555 | 0.07475 | 0.00167 | 0.01152 |
| 80 | 0.57828 | 0.42172 | 21872. | 9224. | 2.42621 | 85619. | 159593. | 7.30 | 0.10773 | 0.10620 | 0.00334 | 0.01679 |
| 85 | 0.0 | 1.00000 | 12648. | 12648. | 5.84865 | 73974. | 73974. | 5.85 | 0.17098 | 0.16850 | 0.00667 | 0.02448 |

FEMALES

| AGE | $_nP_x$ | $_nq_x$ | $l_x$ | $_nd_x$ | $_na_x$ | $_nL_x$ | $T_x$ | $\mathring{e}_x$ | $_nm_x$ | $_nM_x$ | $C_x$ | $B_x$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.90022 | 0.09978 | 100000. | 9978. | 0.31273 | 93143. | 6024905. | 60.25 | 0.10712 | 0.10712 | | |
| 1 | 0.96837 | 0.03163 | 90022. | 2848. | 0.44000 | 88428. | 5931763. | 65.89 | 0.03220 | 0.03220 | | |
| 2 | 0.98634 | 0.01366 | 87175. | 1191. | 0.47000 | 86544. | 5843335. | 67.03 | 0.01376 | 0.01376 | | |
| 3 | 0.99123 | 0.00877 | 85984. | 754. | 0.49000 | 85599. | 5756791. | 66.95 | 0.00881 | 0.00881 | | |
| 4 | 0.99391 | 0.00609 | 85230. | 519. | 0.50000 | 84970. | 5671192. | 66.54 | 0.00611 | 0.00611 | | |
| 5 | 0.98681 | 0.01319 | 84711. | 1118. | 2.49447 | 420754. | 5586222. | 65.94 | 0.00266 | 0.00266 | | |
| 10 | 0.99499 | 0.00501 | 83593. | 419. | 2.22874 | 416806. | 5165469. | 61.79 | 0.00100 | 0.00102 | -0.00001 | -0.00027 |
| 15 | 0.99320 | 0.00680 | 83175. | 566. | 2.60361 | 414517. | 4748662. | 57.09 | 0.00137 | 0.00136 | 0.00001 | 0.00015 |
| 20 | 0.99144 | 0.00856 | 82609. | 707. | 2.56628 | 411323. | 4334146. | 52.47 | 0.00172 | 0.00171 | 0.00000 | 0.00012 |
| 25 | 0.99032 | 0.00968 | 81902. | 793. | 2.61356 | 407617. | 3922822. | 47.90 | 0.00195 | 0.00194 | 0.00001 | 0.00023 |
| 30 | 0.98599 | 0.01401 | 81109. | 1137. | 2.64471 | 402868. | 3515205. | 43.34 | 0.00282 | 0.00281 | 0.00001 | 0.00043 |
| 35 | 0.98026 | 0.01974 | 79972. | 1579. | 2.61119 | 396091. | 3112337. | 38.92 | 0.00399 | 0.00398 | 0.00001 | 0.00048 |
| 40 | 0.97471 | 0.02529 | 78394. | 1983. | 2.60259 | 387215. | 2716246. | 34.65 | 0.00512 | 0.00510 | 0.00002 | 0.00058 |
| 45 | 0.96658 | 0.03342 | 76411. | 2554. | 2.61541 | 375965. | 2329031. | 30.48 | 0.00679 | 0.00676 | 0.00004 | 0.00088 |
| 50 | 0.95411 | 0.04589 | 73857. | 3389. | 2.61660 | 361209. | 1953066. | 26.44 | 0.00938 | 0.00933 | 0.00007 | 0.00128 |
| 55 | 0.93709 | 0.06291 | 70468. | 4433. | 2.62987 | 341834. | 1591858. | 22.59 | 0.01297 | 0.01290 | 0.00009 | 0.00204 |
| 60 | 0.90820 | 0.09180 | 66035. | 6062. | 2.65315 | 315949. | 1250024. | 18.93 | 0.01919 | 0.01912 | 0.00014 | 0.00372 |
| 65 | 0.85603 | 0.14397 | 59973. | 8634. | 2.59893 | 279134. | 934075. | 15.57 | 0.03093 | 0.03077 | 0.00032 | 0.00509 |
| 70 | 0.80256 | 0.19744 | 51339. | 10137. | 2.55777 | 231939. | 654942. | 12.76 | 0.04370 | 0.04354 | 0.00045 | 0.00657 |
| 75 | 0.72776 | 0.27224 | 41202. | 11217. | 2.52094 | 178204. | 423003. | 10.27 | 0.06294 | 0.06230 | 0.00126 | 0.00973 |
| 80 | 0.62919 | 0.37081 | 29985. | 11119. | 2.46287 | 121716. | 244799. | 8.16 | 0.09135 | 0.09026 | 0.00241 | 0.01438 |
| 85 | 0.0 | 1.00000 | 18866. | 18866. | 6.52391 | 123083. | 123083. | 6.52 | 0.15328 | 0.15192 | 0.00462 | 0.02126 |

## IV. METHODOLOGY OF MODEL LIFE-TABLE CONSTRUCTION

This section is adapted from a report by M. Le Bras of the Institut national d'études démographiques, Paris, on work being carried out in consultation with the Population Division. This work is of a methodological nature and involves re-analysis of existing data bases already used for construction of model life-tables and a close scrutiny of the techniques, assumptions and methods used in the construction of existing sets of model life-tables with a view to improving the methodology of summarization of a set of life-tables.

The main methodological questions are the number of parameters to be used; the form of transformation to be used if any; the life-table function to be used; whether some form of clustering is to be used; and the statistical procedures to be used. All of these questions can be addressed at several levels and are inter-linked to some extent.

The first topic is the choice of statistical methodology. Existing model life table systems (except Brass' system) are based on regression techniques. The U.N. tables were calculated by 'chained' regression procedures with ${}_1q_0$ taken as the independent variable in regressions of ${}_nq_x$'s for separate age intervals. Such a procedure is efficient in producing model life-tables from a starting point of ${}_1q_0$ (ignoring the possible problems of the 'chaining' procedure). In practice, users of model life-tables are very unlikely to have only an estimate of ${}_1q_0$ and will usually not have such an estimate at all. Using some other entry indicator effectively requires inversion of the equations and efficiency (and accuracy) is lost. It is well-known that the regression of one variable on another does not give the same relationship as if the variables are reversed, owing to the nature of the error assumptions implicit in the regression model. Thus the U.N. (1955, 1956) tables are most efficient when the entry actually used is ${}_1q_0$. Similarly Coale and Demeny (1966) calculated all of their regressions in terms of $\mathring{e}_{10}$. Whilst they justify this as a good overall indicator of level of mortality, there is still an inversion problem when

other measures are used as entries to their tables. In practice the equations are not actually inverted, but resort is made to published tables. Nevertheless statistical efficiency is necessarily lost. The Ledermann (1969) tables recognized these problems and set out to construct tables in terms of various explicit entry values, which were chosen as likely to occur in practice. However, such an approach presents problems. A different set of model life-tables is required for every possible set of inputs for real efficiency. Thus, a major problem with a regression-based approach is the lack of a symmetrical treatment of all the variables. For many uses of model life-tables (e.g., projections, smoothing a series of age-specific rates, comparisons with a series of intercensal survivorship ratios), there is not a single entry point, or even two or three. Another difficulty which can arise is that regression based on, say, two entry values may be very inefficient at summarising variability in mortality experience at some other ages, even though giving the most efficient prediction from the two input variables.

For many of the potential uses of model life-tables (especially as a reference set), we are most interested in obtaining an efficient summary of the input data set. This is a data analysis rather than a statistical inference problem. The most common approach to such summary data analysis would be through principal components analysis. Such an analysis is more efficient than regression at summarizing overall variability within a data set and avoids taking any single variable as a predictor. In other words, principal components treats all the variables symmetrically. Thus, although a case can be made for regression models to provide the most efficient predictors of mortality at other ages from a small number of items of information on mortality, it is felt that the very broad range of uses to which any set of model life-tables are likely to be applied strongly indicates the use of a summarisation or data analytic approach, through principal components analysis.

To demonstrate and illustrate the advantages of principal components analysis, a series of comparisons were made. For the 193 observed life-tables

used by Coale and Demeny in the actual derivation of their regional model life-tables, the amounts of residual variation were considered from each of three fits. For the Coale-Demeny models the fit was made by choosing the region to which the table had been assigned and taking the model life-table within that regional set with the same $\overset{\circ}{e}_{10}$, thus taking a pure approach to the methodology of regression used. For the Ledermann fit, a regression analysis was carried out relating $\ln {}_nq_x$ to $\ln {}_5q_0$ and $\ln {}_{20}q_{45}$ as the entries. These values within the observed life-tables (${}_5q_0$ and ${}_{20}q_{45}$) were then used to generate the fitted table. Finally, a principal components analysis was carried out on the correlation matrix of $\ln {}_nq_x$'s and the resulting models fitted to each of the observed tables in turn. The models used were based on two components and are described later. To avoid placing too much trust in a single distance measure, the residuals of each of the sets of fitted tables from the set of observed tables were treated in several ways, using as measures of distance between the observed $q_x^i$ and the fitted $\hat{q}_x^i$ the following:

$$S_1 = \sum_x \sum_i \left( q_x^i - \hat{q}_x^i \right)^2$$

$$S_2 = \sum_x \sum_i \left| q_x^i - \hat{q}_x^i \right|$$

$$S_3 = \sum_i \max_x \left\{ \left| q_x^i - \hat{q}_x^i \right| \right\}$$

$$S_4 = \sum_x \sum_i \left( \ln q_x^i - \ln \hat{q}_n^i \right)^2$$

$$S_5 = \sum_x \sum_i \left| \ln q_x^i - \ln \hat{q}_x^i \right|$$

$$S_6 = \sum_i \max_x \left\{ \left| \ln q_x^i - \ln \hat{q}_x^i \right| \right\}$$

Table IV.1 Distance between observed and fitted tables for model life-table systems

| Distance measure* | Coale and Demeny | Lederman | Principal components |
|---|---|---|---|
| | F e m a l e s | | |
| $S_1$ (Euclidean) | .258 | .402 | .216 |
| $S_2$ (absolute value) | 16.0 | 20.6 | 14.1 |
| $S_3$ (maximum) | 4.42 | 5.13 | 3.77 |
| $S_4$ (As above but using logarithms) | 292.2 | 119.7 | 68.1 |
| $S_5$ | 411.0 | 448.4 | 298.6 |
| $S_6$ | 117.7 | 72.1 | 67.1 |
| | M a l e s | | |
| $S_1$ | .407 | .420 | .355 |
| $S_2$ | 21.3 | 21.4 | 16.9 |
| $S_3$ | 5.34 | 5.32 | 5.30 |
| $S_4$ | 385.3 | 83.4 | 52.2 |
| $S_5$ | 485.1 | 367.7 | 269.6 |
| $S_6$ | 119.0 | 62.3 | 61.2 |

* See text for definition of distance measures.

Table IV.1 shows the values obtained for these various distance measures for each of the sets of model life-tables separately for each sex. Whichever distance measure is used, principal components provides a better summary of the 193 observed tables. It was almost certain that principal components would do better than the Ledermann regressions, especially on $S_4$ because of its minimisation of virtually the same distance measure. But it was not obvious in advance that this approach through principal components would outperform the Coale and Demeny approach, with its identification of supposedly homogeneous subsets of the data. In addition to these overall distance measures, pairwise comparisons were carried out on each of the distance measures to check that the differences did not arise through a few aberrant observations. Table IV.2 shows the results of these pairwise comparisons of both the Coale and Demeny and the Ledermann models with the set based on principal components. Again it is clear that the principal components model does better than either the Coale and Demeny models or the Ledermann approach in terms of goodness-of-fit to all of the original data set. These findings reinforce quite strongly the theoretical arguments for use of principal components, rather than regression. Nevertheless it is still the case, for example, that given only ${}_5q_0$ and ${}_{20}q_{45}$ the Ledermann equations would better predict the other life-table values than would the principal components model. There is always such a trade-off between the generality of the principal components approach and the particularity of a regression-based approach. It is of some interest to see where the differences between the systems arise, and

Table IV.2. Numbers of observed tables for which the fit is better on a pairwise comparison of model life-table systems, various distance measures.

| Distance measure | Principal components | | Principal components | |
|---|---|---|---|---|
| | Better than Coale and Demeny | Worse than Coale and Demeny | Better than Ledermann | Worse than Ledermann |
| | F e m a l e s | | | |
| $S_1$ | 132 | 61 | 140 | 53 |
| $S_2$ | 133 | 60 | 154 | 39 |
| $S_3$ | 121 | 72 | 127 | 66 |
| $S_4$ | 121 | 72 | 147 | 46 |
| $S_5$ | 118 | 75 | 148 | 45 |
| $S_6$ | 122 | 71 | 118 | 75 |
| | M a l e s | | | |
| $S_1$ | 121 | 72 | 121 | 72 |
| $S_2$ | 135 | 58 | 146 | 47 |
| $S_3$ | 112 | 81 | 106 | 87 |
| $S_4$ | 125 | 68 | 135 | 58 |
| $S_5$ | 130 | 63 | 145 | 48 |
| $S_6$ | 117 | 76 | 97 | 96 |

as an example Table IV.3 shows the number of tables (for males) better represented by the principal components and Coale and Demeny models by the region to which they belong (using $S_4$ as the distance measure). It is clear that the Coale-Demeny models perform better within the smaller subsets of

Table IV.3. Comparison of fits to observed tables by using $S_4$ for males.

| | Region | | | | |
|---|---|---|---|---|---|
| | West | North | East | South | Total |
| Coale Demeny better | 30 | 7 | 16 | 16 | 68 |
| Principal components better | 101 | 2 | 15 | 6 | 125 |

the observed life-tables. This should not really be surprising, as choosing 9 life-tables for their homogeneity and then fitting each of 17 regressions to 9 data points should give a close fit. For the larger sets, especially West, the advantage of a second parameter becomes apparent and the principal components model fits better.

Some further results were obtained using the 326 tables originally considered by Coale and Demeny before exclusion of certain tables. The results of comparisons between the three approaches were similar to those above, although the relative advantage of principal components became greater if anything and all the residual mean square statistics (on either natural or logarithmic scales) became larger, indicating a noticeable worsening of overall fit. However, these calculations resulted in some interesting

Table IV.4. Re-assignment of Coale and Demeny tables to regions

| Region assigned by Coale and Demeny | Region giving best fit | | | | | |
|---|---|---|---|---|---|---|
| | West | North | East | South | Unspecified | Total |
| | | | F E M A L E S | | | |
| West | 79 | 8 | 10 | 0 | 34 | 131 |
| North | 0 | 7 | 0 | 0 | 2 | 9 |
| East | 0 | 0 | 28 | 1 | 2 | 31 |
| South | 0 | 0 | 0 | 19 | 3 | 22 |
| Other tables | 33 | 21 | 15 | 3 | 61 | 133 |
| Total | 112 | 36 | 53 | 23 | 102 | 326 |

| Region assigned by Coale and Demeny | Region giving best fit | | | | | |
|---|---|---|---|---|---|---|
| | West | North | East | South | Unspecified | Total |
| | | | M A L E S | | | |
| West | 63 | 8 | 14 | 5 | 41 | 131 |
| North | 1 | 7 | 0 | 0 | 1 | 9 |
| East | 3 | 0 | 24 | 1 | 3 | 31 |
| South | 1 | 0 | 0 | 20 | 1 | 22 |
| Other tables | 38 | 33 | 8 | 3 | 51 | 133 |
| Total | 106 | 48 | 46 | 29 | 97 | 326 |

findings. One of the tasks that was carried out was the allocation of the 133 life-tables excluded by Coale and Demeny to their 'nearest' region. In addition and using the same fitting criteria, each of the original 193 tables were allocated to their 'nearest' region. (The metrics used were $S_1$, $S_2$ and $S_3$ and allocation was 'unspecified' unless all three criteria indicated the same region). The results of this allocation were surprising and are shown at Table IV.4. Of the 193 tables assigned to regions by Coale and Demeny, 19 (for females) and 33 (for males) were unambiguously better approximated by a model life-table from a different region and a further 41 (female) and 46 (male) were not clearly allocable to a region as they were inconsistent on the three fitting criteria. It should be noted at this point that these fits were carried out in terms of $\delta_{10}$, the Coale and Demeny regression variable, rather than any measure of overall fit. Coale and Demeny made their regional allocations on the basis of profiles of deviations. However, such results raise questions about the identification of families of model life tables. It is this question that will be considered next.

The Coale and Demeny profiles of mortality deviation, which were the basis for their classification procedures, have been subjected to substantial re-analysis through automated clustering and classification procedures. In general these analyses suggest that pattern clusters as identified by Coale and Demeny could equally well be regarded as picking out extreme (or moderately extreme) parts of a continuous distribution. Although it is difficult to demonstrate these findings compactly, an attempt to do so is provided in Figure IV.5. This chart shows the positions of the Coale-Demeny male data set in the plane of the first two principal components of the profiles of deviation (i.e. after controlling for overall level of mortality). About 75% of the total variability is accounted for by these two components. The regional tables are indicated by their initial letters (N, E, S and W), with the remaining tables being indicated by reference numbers. Whilst the tables allocated by Coale and Demeny to particular regions are somewhat clustered, it is at least as plausible to argue that these are just areas of a continuous bivariate distribution of profile patterns. As one quarter of

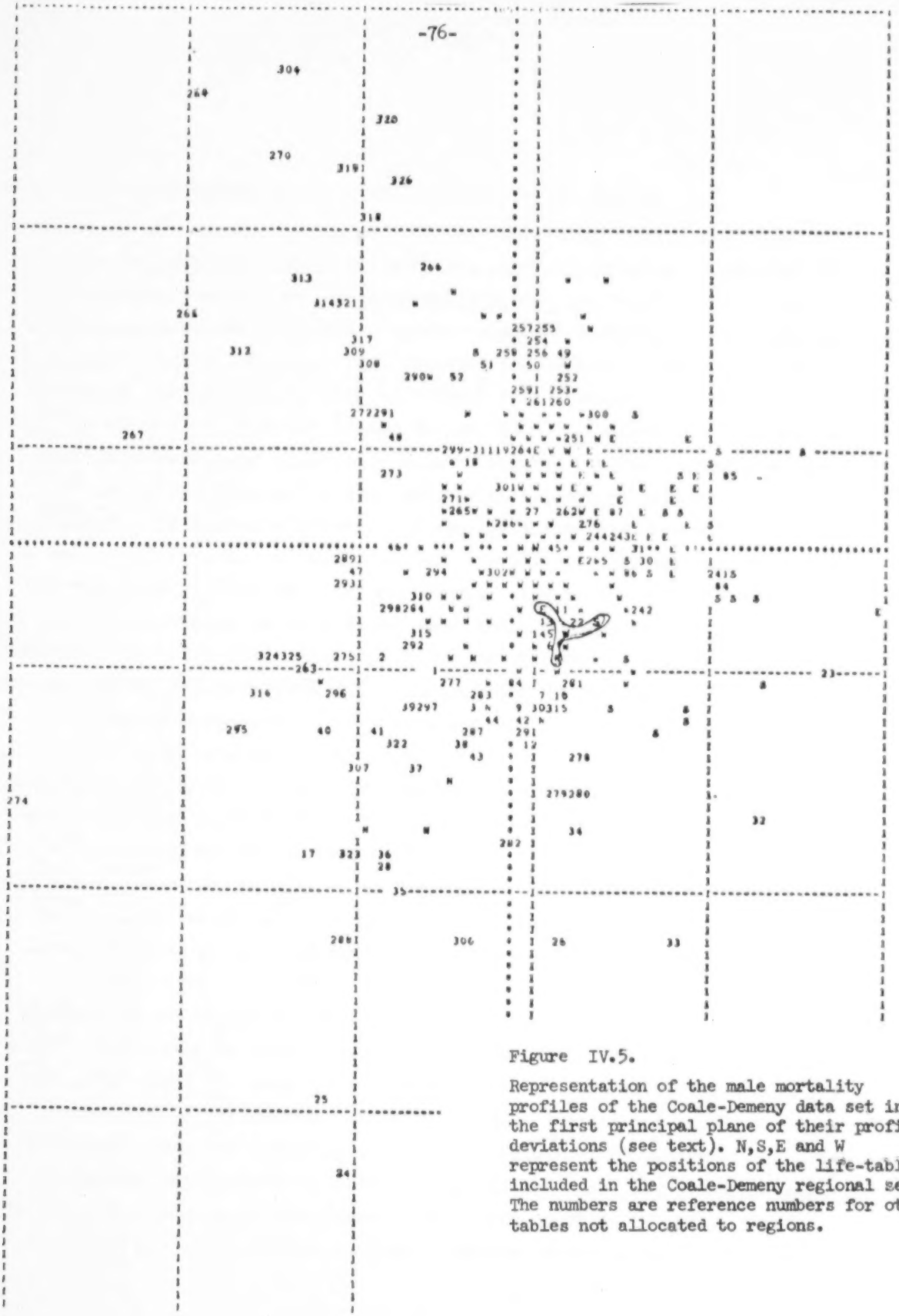


Figure IV.5.

Representation of the male mortality profiles of the Coale-Demeny data set in the first principal plane of their profile deviations (see text). N,S,E and W represent the positions of the life-tables included in the Coale-Demeny regional sets. The numbers are reference numbers for other tables not allocated to regions.

the variability was not accounted for by the first two principal components too much should not be made of Figures IV.5 and 6. However, application of automatic clustering algorithms confirms this finding, namely that it is preferable to identify extra dimensions of mortality, rather than trying to identify subsets. Of course, a further limitation of automatic clustering techniques is that they take no account of geographic and temporal proximity of tables. It is a matter of debate as to whether this is correct as geographical clustering has intuitive appeal as well as practical value. At this stage and subject to classificatory work on further input sets, it seems that identification of an extra dimension or more is preferable to a clustering approach.

It is now necessary to consider some further crucial problems in the methodology of construction of model life-tables, namely the number of dimensions or parameters to be used, the life-table function to be used and transformations to be considered. All of these are closely interlinked. The comparative work so far has been carried out for the following functions and transformations:

${}_nq_x$ , ${}_nd_x$ and $l_x$; $\ln({}_nq_x)$; and logit $(l_x)$ $\left( = {}^1/_2 \ln \left( \frac{l_o - l_x}{l_x} \right) \right)$. Other transformations may also be tried in subsequent work.

The first set of results relates to the dimensionality of mortality (the number of parameters necessary). This work has been carried out on both the 147 tables used by the United Nations to construct their tables and the 326 tables of Coale and Demeny (not just the 193 tables used for the construction of the regional tables, which were more homogeneous). The findings were similar for both data sets and for each sex. The results are presented at Table IV.7. All of these principal components analyses were carried out using correlations rather than covariances (in other words the variables were standardized to zero mean and unit variance). This is more justifiable with logarithmic tranforms (i.e. $\ln({}_nq_x)$ and logit $(l_x)$) than with untransformed variables, where covariances are probably better, in the sense

of being sensitive to high variability and thus providing a better fit in terms of residual variance. In fact, by the very nature of the distance measures used, covariance analysis will always lead to the minimum residual variance on the metric used; in our earlier notation $S_1$ is minimized by covariance analysis for untransformed ${}_nq_x$'s, and $S_4$ is minimized for $\ln({}_nq_x)$'s. Thus, for best overall fit with untransformed data the covariances should be used. However, at least in this case, a principal components analysis in terms of $\ln({}_nq_x)$ better represents the ${}_nq_x$'s if carried out in terms of correlations than in terms of covariances.

As a result of these considerations, Table IV.7 does not give the best fitting components for the untransformed variables, which would be based on covariance analysis. Although this is in process of being revised, the results at Table IV.7 are the only ones available at the time of writing. Turning now to the interpretation of Table IV.7, it is clear that a fairly good to good representation can always be obtained with two dimensions and a very good one with three or four dimensions. It must be noted that the percentages shown in Table IV.7 are percentages of total variability of the particular function used and as a result are not strictly comparable. Even though very high proportions of variance are accounted for with $l_x$ or logit $(l_x)$, this does not necessarily mean that ${}_nq_x$ will be better represented by the $l_x$ based tables, for example.

It is this problem of mutual representation of one function by an analysis based on another that is next considered. At this stage the analysis was based on models with only the first two principal components, again with correlations rather than covariances. The analysis in terms of logits was carried out as R-analysis rather than Q-analysis, which is the dual of usual principal components analysis, taking the tables as variables and the ages as the observations. No good data analytic reason is known for doing this, but it is more in line with the spirit of Brass' (1971) work on logits. Again both the 147 U.N. tables and the 326 Coale and Demeny tables were used. For each method of fitting, the two principal components were derived; each observed table was fitted (by ordinary least squares) in terms of the input

Table IV.7. Percentages of variance explained by successive principal components

| Tables | Variable | Sex | Component 1st | 2nd | 3rd | 4th | Total | 1st 2 |
|---|---|---|---|---|---|---|---|---|
| | $_nq_x$ | M | 80.2 | 8.5 | 5.2 | 2.3 | 96.2 | 88.7 |
| | | F | 84.1 | 6.4 | 3.8 | 1.9 | 96.2 | 90.5 |
| UNITED | ln ($_nq_x$) | M | 81.0 | 10.5 | 3.8 | 2.0 | 97.3 | 91.5 |
| | | F | 84.6 | 8.0 | 3.6 | 1.7 | 97.9 | 92.6 |
| NATIONS | $l_x$ | M | 94.1 | 4.2 | 1.2 | 0.3 | 99.8 | 98.3 |
| (147 | | F | 95.9 | 2.6 | 1.0 | 0.3 | 99.8 | 98.5 |
| tables) | $_nd_x$ | M | 68.0 | 19.4 | 5.1 | 2.9 | 95.4 | 87.4 |
| | | F | 71.3 | 17.8 | 4.2 | 2.6 | 95.9 | 89.1 |
| | logit ($l_x$) | M | 99.5 | 0.3 | 0.2 | - | 100.0 | 99.8 |
| | | F | 99.4 | 0.3 | 0.2 | - | (99.9) | 99.7 |
| | $_nq_x$ | M | 82.3 | 7.8 | 3.6 | 1.9 | 95.6 | 90.1 |
| COALE | | F | 82.5 | 4.3 | 3.5 | 3.0 | 93.3 | 86.6 |
| and | ln ($_nq_x$) | M | 85.5 | 8.5 | 2.7 | 1.0 | 97.7 | 94.0 |
| DEMENY | | F | 90.4 | 4.2 | 2.3 | 0.9 | 97.8 | 94.6 |
| (326 | $l_x$ | M | 96.3 | 2.7 | 0.7 | 0.2 | 99.9 | 99.0 |
| | | F | 97.2 | 1.8 | 0.6 | 0.2 | 99.8 | 99.0 |
| tables) | $_nd_x$ | M | 71.2 | 18.9 | 3.0 | 2.8 | 95.9 | 90.1 |
| | | F | 73.0 | 18.5 | 2.8 | 1.8 | 96.1 | 91.5 |
| | logit ($l_x$) | M | 99.4 | 0.3 | 0.2 | - | (99.9) | 99.7 |
| | | F | 99.1 | 0.6 | 0.2 | - | (99.9) | 99.7 |

variable; and the deviations at each age were calculated not only for the generating variable, but for the others as well and the sums of squares of these deviations were calculated. The results of this very important analysis are presented at Table IV.8. It is obvious from even the most cursory inspection of the columns of Table IV.8 that the models based on $_nd_x$ perform

Table IV.8. Comparison of squared errors

| Principal components calculated using | Residuals for function | | | | |
|---|---|---|---|---|---|
| | $_nq_x$ | $\ln(_nq_x)$ | $l_x$ | $_nd_x$ | logit $(l_x)$ |
| Sum of squared residuals with principal components across ages and countries (both sexes combined) based on 148 tables of the United Nations: | | | | | |
| $_nq_x$ | 0.516 | 82.71 | 2.78 | 0.281 | 29.32 |
| $\ln(_nq_x)$ | 0.712 | 59.62 | 3.32 | 0.336 | 26.65 |
| $l_x$ | 2.786 | 244.09 | 0.461 | 0.147 | 47.36 |
| $_nd_x$ | 7.604 | 834.86 | 3.521 | 0.248 | 157.91 |
| logit $(l_x)$ | 1.054 | 101.87 | 0.586 | 0.287 | 6.00 |
| based on 326 tables of Coale and Demeny: | | | | | |
| $_nq_x$ | 1.286 | 355.4 | 3.52 | 0.559 | 44.83 |
| $\ln(_nq_x)$ | 1.236 | 128.2 | 5.56 | 0.662 | 50.36 |
| $l_x$ | 4.470 | 836.7 | 0.711 | 0.328 | 80.72 |

very badly: this is undoubtedly partly related to the use of correlation rather than covariance analysis, as the numbers of deaths vary very substantially with age, as does their variance. It is still fairly certain that an analysis based on the deaths function would not perform well overall, even if based on covariance analysis. This leaves comparisons of the remaining four approaches. At this stage, it becomes important to decide which life-table functions need to be well represented. For example, it is probably of little importance that either logit ($l_x$) or ln ($_nq_x$) is well represented in view of the uses to which model life-tables are put (perhaps with the minor exceptions of their use as reference standards within Brass' logit transformation system and the use of semi-logarithmic plots of $_nq_x$). To improve the comparisons of the various approaches it is necessary to remove the differences in magnitude of the various measures of error. This has been attempted by scaling each column so that it averaged 1.0 (with the deaths-based fits excluded). Table IV.9 shows the results of this scaling operation. From this table it would seem that the best all-round compromise would be to use the logit-based analysis, which performs fairly well at representing all the variables and avoids doing very badly on any. It is perhaps surprising that the logit-based analysis represents $l_x$ quite well but the reverse is not the case. The fits obtained with analysis on $_nq_x$ are generally better than those on ln ($_nq_x$), although several previous studies have preferred to use $\ln(_nq_x)$ as the basis of model life-tables (most notably Ledermann who chose $\ln(_nq_x)$ to reduce heteroscedasticity, although it is doubtful whether such considerations are important for a data-analytic approach concerned only with goodness-of-fit, being much more related to the assumptions necessary for inference). These results are not yet conclusive and logit ($_nq_x$) also deserves investigation, as do covariance-based approaches for untransformed data. Although at an early stage, such methodological studies are clearly useful and important.

Table IV.9. Scaled residual sums of squares

| Tables | Scaled residuals for function | Principal components fitted through: | | | |
|---|---|---|---|---|---|
| | | $_nq_x$ | ln ($_nq_x$) | $l_x$ | logit ($l_x$) |
| UNITED NATIONS | $_nq_x$ | .407 | .562 | 2.199 | .832 |
| | ln ($_nq_x$) | .678 | .488 | 1.999 | .835 |
| | $l_x$ | 1.556 | 1.858 | .258 | .328 |
| | logit ($l_x$) | 1.073 | .975 | 1.733 | .220 |
| | $_nd_x$ | 1.069 | 1.279 | .559 | 1.092 |
| COALE and DEMENY | $_nq_x$ | .574 | .552 | 1.996 | .878 |
| | ln ($_nq_x$) | .902 | .325 | 2.123 | .649 |
| | $l_x$ | 1.315 | 2.042 | .261 | .381 |
| | logit ($l_x$) | .959 | 1.078 | 1.727 | .236 |
| | $_nd_x$ | 1.011 | 1.197 | .593 | 1.199 |
| Total | $_nq_x$, $l_x$ and $_nd_x$ | 5.932 | 7.490 | 5.866 | 4.710 |
| Total | All | 9.544 | 10.356 | 13.448 | 6.650 |

This set of model life-tables was derived during the analysis of the appropriate statistical methodology. From examination of Tables IV.8 and IV.9, it is clear that the use of $\ln ({}_nq_x)$ was probably not the best choice, although it is fairly efficient at summarising variability in ${}_nq_x$. It is clear from the analysis shown in Table IV.1 and IV.2 that these model tables do better than existing model life-table sets in terms of overall fit to observed data sets (or at least the set of tables used by Coale and Demeny). For our initial purposes, namely to fit model life-tables to ${}_nq_x$ values, they are almost as efficient as any.

These were the models used for the analyses discussed in Section III.6-b. The principal components model can be written as $\ln ({}_nq_x)^i = U_{xo} + a_i U_{x1} + b_i U_{x2}$ in this case, where the $U_{xj}$'s are the principal components and $a_i$ and $b_i$ are the two parameters associated with the i th table. Table IV.10 shows the values of the $U_{xj}$'s (normalised so that $\sum_x U_{xj}^2 = 1$ for all j) and the proportions of variance explained by the first and first two components for each age-group. The fit is least good for the earliest and latest ages, especially for males. Some of this lack of fit at the extreme ages is no doubt due to errors, as the use of the entire collection of 326 tables avoids any screening and rejection of erroneous tables, which is a necessary step. Nevertheless, it is useful at this stage to have this preliminary set of principal components based model life-tables available, even though a better set may yet be found.

To conclude, many of the results of this methodological research are already clear. Others need further study, but a brief summary of the main points is in order. The work suggests the use of principal components except for very specific purposes; in general model life-tables are put to varied uses and principal components will be the most satisfactory approach. Two or three parameters are probably adequate to represent variation in age-patterns of mortality. Certain functions can be rejected at this stage as a basis for deriving the model life tables (e.g. ${}_nd_x$). There is some

Table IV.10 Principal components and proportions of variances explained for model tables based on $\ln {}_nq_x$ for Coale and Demeny data set.

| x | $U_{x0}$ | $U_{x1}$ | $U_{x2}$ | Percentage of variance explained | |
|---|---|---|---|---|---|
| | | | | 1st component | 1st and 2nd component |
| | | | Males | | |
| 1 | -2.34336 | .24950 | -.21742 | 76.97477 | 81.65995 |
| 2 | -3.22588 | .41952 | -.51031 | 81.97884 | 91.70092 |
| 3 | -4.27702 | .34105 | -.45369 | 82.66976 | 94.39467 |
| 4 | -4.61872 | .31041 | -.37154 | 83.50521 | 93.09345 |
| 5 | -4.16474 | .26845 | -.27780 | 87.26640 | 94.75663 |
| 6 | -3.80664 | .27173 | -.26414 | 88.67003 | 95.38518 |
| 7 | -3.74907 | .28348 | -.21081 | 92.78342 | 96.89574 |
| 8 | -3.63621 | .28074 | -.15907 | 94.86172 | 97.30258 |
| 9 | -3.45750 | .26932 | -.08339 | 96.33319 | 97.07335 |
| 10 | -3.19618 | .24016 | -.01303 | 96.68565 | 96.70845 |
| 11 | -2.91198 | .19976 | .07233 | 94.54171 | 95.53512 |
| 12 | -2.57749 | .16540 | .12421 | 90.86302 | 94.96990 |
| 13 | -2.25006 | .13351 | .17274 | 83.97838 | 95.24585 |
| 14 | -1.88823 | .11340 | .16113 | 83.27457 | 96.74819 |
| 15 | -1.53369 | .09033 | .15176 | 78.57035 | 96.34655 |
| 16 | -1.17428 | .07199 | .12382 | 77.29411 | 95.61944 |
| 17 | - .83813 | .05008 | .09903 | 58.99402 | 77.48098 |
| | | | Females | | |
| 1 | -2.52154 | .23127 | -.07236 | 83.33939 | 83.60865 |
| 2 | -3.29097 | .39007 | -.39561 | 87.56313 | 90.53563 |
| 3 | -4.37976 | .35072 | -.34354 | 90.81929 | 93.69501 |
| 4 | -4.69823 | .32742 | -.36685 | 91.26520 | 95.04625 |
| 5 | -4.28856 | .31232 | -.34937 | 92.86849 | 96.70385 |
| 6 | -3.99748 | .30228 | -.35774 | 92.02004 | 96.27354 |
| 7 | -3.85390 | .29044 | -.30584 | 94.50954 | 97.96820 |
| 8 | -3.72941 | .27373 | -.23887 | 96.18266 | 98.59981 |
| 9 | -3.57980 | .24776 | -.16807 | 96.96048 | 98.43298 |
| 10 | -3.40719 | .21518 | -.09373 | 96.51398 | 97.11827 |
| 11 | -3.19096 | .17846 | .00281 | 93.93007 | 93.93084 |
| 12 | -2.88191 | .15724 | .06742 | 92.69968 | 93.26210 |
| 13 | -2.56737 | .13596 | .14142 | 84.04169 | 87.04257 |
| 14 | -2.15279 | .12178 | .15775 | 90.39624 | 95.40222 |
| 15 | -1.74913 | .09875 | .18327 | 87.67484 | 97.64219 |
| 16 | -1.32797 | .07645 | .17661 | 82.04200 | 96.49365 |
| 17 | - .94478 | .05266 | .18377 | 60.79128 | 85.22457 |

The problems of how to link the sexes are under consideration but not yet at a stage where anything can be reported, although some progress is hoped for through canonical correlation analysis.

## CONCLUSION

At this stage of the project many crucial decisions have already been made. The project was early on committed to making maximal use of data from developing countries, which made it essential to adopt extensive evaluation and adjustment procedures. While every effort has been made to use as many such procedures as were routinisable, others may still be incorporated or developed, but within the limitations of routinisation through computer programming. Several of the procedures tried have proven less satisfactory than was initially hoped, especially the Hogan procedure on mortality under age 15, but also the adaptation of Martin's technique. Inevitably, with such a battery of assessment procedures, not all of which always give consistent results, any final adjustments adopted are somewhat subjective. This kind of analysis and adjustment is more of an art than an exact science. It is for these reasons that it is essential to provide an adequate, detailed record of such analyses to permit others to assess the results. An idea which is still being explored is provision of subjectively chosen weights for the adjusted life-tables to reflect judgement of their relative accuracy. Thus, for El Salvador higher weight might be given to the 1971 adjusted table, as the childhood mortality estimates are more firmly based therein.

Although every attempt is being made to produce adjusted life-tables which are as accurate measures as possible of the mortality conditions and patterns of the countries concerned, there must remain reservations. Until either data collection is improved dramatically (by which time mortality will have changed, perhaps substantially) or fundamentally new and powerful assessment and adjustment techniques appear, these doubts will remain. Nevertheless these carefully adjusted life-tables should prove invaluable as improved estimates of mortality for the countries concerned, and they will be as good inputs as can be found from developing countries at this time. As such, model life-tables based on them will be a valuable alternative to existing sets, with some of the doubts about accuracy and validity being counter-

Once the adjusted life-tables are available, the remaining problems of generation of the new model tables have to be faced. This stage seems fairly certain to be approached through principal components analysis, with use of two parameters appearing the most likely outcome, although this problem of dimensionality will be solved by examining the data at hand. It is not yet certain which life-table function and transformation (if any) will be used; logit ($l_x$) or logit ($_nq_x$) appear suitable, but $_nq_x$ also has attractions. Depending on whether transformations are used or not, the choice between correlation or covariance analysis has to be made: without transformations covariances are probably preferable. Despite the questions raised earlier about the regional patterns of Coale and Demeny, the tables will be examined for clustering, although the number of tables finally available may not be large enough to allow distinction of families.

Some further work is necessary. A technique for handling interpolation and comparison of mortality at the young ages would still be invaluable. For all of the growth-balance techniques it would be extremely useful if a review could be made of studies of completeness of death registration (like EDENH for example) in the developing world, with a view to refining the assumptions of constant underregistration with age, perhaps even leading to a parameterization of such patterns. Such a parameterization, even if simple, would make the techniques harder to apply, requiring iterative non-linear fitting procedures. Even so, such extra effort could prove extremely productive, especially as such fitting procedures can be routinized.

A further area, which is beyond the scope of this project but warrants considerable further investigation, is the whole aspect of utilization of cause of death data in building model life-tables. Some progress could undoubtedly be made by studying typical levels and variability in cause-specific rates, controlling for overall level, and utilizing mixes of cause-specific rates to generate model life-tables. For example, in view of the known high incidence of diarrhoea and gastro-enteritis in many developing countries, tables could be generated with high values on this component. Relevant to such work, but by

REFERENCES

G. Barclay, et al. (1976) "A Reassessment of the Demography of Traditional Rural China" Population Index 42(4):606-635.

M.S. Bartlett (1949) "Fitting a Straight Line when Both Variables are Subject to Error" Biometrics 5:207-212.

B. Benjamin and H.W. Haycocks (1970) The Analysis of Mortality and Other Actuarial Statistics Cambridge University Press.

L.R. Bixby (1976) El Sistema Modelo de Brass: El Salvador, 1961-1971. CELADE, Series C, No. 1001.

J. Bourgeois-Pichat (1946) "De la Mesure de la Mortalité Infantile" Population 1:53-68.

W. Brass (1971) "A Critique of Methods for Estimating Population Growth in Countries with Limited Data" Bulletin of the International Statistical Institute 44(1):397-412.

W. Brass (1975) Methods of Estimating Fertility and Mortality from Limited and Defective Data Laboratories for Population Statistics, Chapel Hill, North Carolina.

P. Cantrelle (1976) "Is There a Standard Pattern of Tropical Mortality" in P. Cantrelle (Ed.) Population in African Development. I.U.S.S.P.

P. Cantrelle and H. Leridon (1971) "Breast Feeding, Mortality in Childhood and Fertility in a Rural Zone of Senegal" Population Studies 25(3): 505-533.

A.J. Coale and P. Demeny (1966) Regional Model Life-Tables and Stable Populations Princeton University Press.

A.J. Coale and N. Goldman (1978) "Notes on a Far-Eastern Pattern of Mortality" Princeton University (mimeographed).

J. Durbin (1954) "Errors in Variables" Review of International Statistical Institute 22:23-32.

T. Dyson (1977) "Levels, Trends, Differentials and Causes of Child Mortality - A Survey" World Health Statistics Report 30(4):282-311.

J.E. Gordon, J.B. Wyon and W. Ascoli (1967) "The Second Year Death Rate in Less Developed Countries" American Journal of Medical Sciences 254: 357-380.

REFERENCES (continued)

F.S. Harper (1936) "An Actuarial Study of Infant Mortality" Skandinavisk Aktuarietidskrift 19:234-270.

H.R. Hogan (1976) Age Patterns of Infant Mortality Unpublished Ph.D. Thesis, Princeton.

M.H.M. Husein (1978) The Age Pattern of Mortality as a Function of Cause of Death Structure Unpublished Ph.D. Thesis, University of London.

N. Keyfitz (1977) Applied Mathematical Demography Wiley and Sons.

N. Keyfitz and J. Frauenthal (1975) "An Improved Life Table Method" Biometrics 31:889-899.

S. Ledermann (1969) Nouvelles Tables-Types de Mortalité (Cahier No. 53 de l'Institut National d'etudes Demographiques). Presses Universitaires de France.

R. Lundy (1978) Age Patterns of Mortality Unpublished Ph.D. Thesis, Berkeley, University of California.

L. Martin (1978) "Completeness of Death Registration: An Evaluation of the Brass Technique" Paper delivered at Population Association of America meeting, Atlanta, April 12/15, 1978. (mimeographed).

L. Martin (1977) Measuring Completeness of Death Registration in Destabilized Populations. Unpublished Ph.D. dissertation, Princeton University.

D.R. McNeil (1977) Interactive Data Analysis Wiley and Sons, New York.

F. Mosteller and J.W. Tukey (1977) Data Analysis and Regression Addison-Wesley.

S. Oppermann (1870) The Insurance Record Feb. 11, 1870.

Organisation for Economic Cooperation and Development (1979) Mortality Project: Annotated Bibliography on the Sources of Demographic Data Vol. I: Africa and the Middle East; Vol. II: Latin America and the Caribbean; Vol. III: Asia. OECD

A. Ortega and M. Rincón (1975) Encuesta Demográfica Nacional de Honduras, Fasciculo IV, Mortalidad. Centro Latinoamericano de Demografia (CELADE) Series A, No. 129.

L. Polissar (1973) Parameterizing Age Distributions of Death by Cause Unpublished Ph.D. Thesis, Princeton.

REFERENCES (continued)

J. Potter and M. Ordóñez G. (1976) "The Completeness of Enumeration in the 1973 Census of the Population of Colombia" Population Index 42(3): 377-403.

S.H. Preston (1978) "Estimating the Completeness of Death Registration" (Mimeographed) United Nations, Population Division.

S.H. Preston (1976) Mortality Patterns in National Populations: With Special Reference to Recorded Causes of Death Academic Press.

S.H. Preston, N. Keyfitz and R. Schoen (1972) Causes of Death: Life Tables for National Populations New York: Seminar Press.

H.S. Shryock and J.S. Siegel (1971) The Methods and Materials of Demography (2 vols.) U.S. Bureau of Census, Government Printing Office.

G.B. Simmons, et al. (1978) Patterns and Causes of Infant Mortality in Rural Uttar Pradesh (Mimeographed).

A.S. Soliz Sánchez (1976) América Central: Estimaciones de la mortalidad infantil y juvenil CELADE, Series C, No. 1003.

J.F. Steffensen (1930) "Infantile Mortality from an Actuarial Point of View" Skandinavisk Aktuarietidskrift 13:272-286.

J.M. Sullivan (1973) "The Influence of Cause-Specific Mortality Conditions on the Age Pattern of Mortality with Special Reference to Taiwan" Population Studies 27:135-158.

United Nations (1955) Age and Sex Patterns of Mortality: Model Life-Tables for Underdeveloped Countries (Population Studies No. 22) New York.

United Nations (1956) Manual II: Methods of Appraisal of Quality of Basic Data for Population Estimates ST/SOA(Series A/23). New York.

United Nations (1956) Manual III: Methods for Population Projections by Sex and Age (Population Studies No. 25) New York.

United Nations (1977) World Population Prospects as Assessed in 1973 New York.